2013

A Survey on Inter-Cell Interference Coordination Techniques in OFDMA-Based Cellular Networks

Abdelbaset S. Hamza
University of Nebraska-Lincoln, ahamza@cse.unl.edu

Shady S. Khalifa
Cairo University, Egypt

Haitham S. Hamza
Department of Information Technology, Cairo University, Egypt

Khaled Elsayed
Department of Electronics and Communications, Cairo University, Egypt

A Survey on Inter-Cell Interference Coordination Techniques in OFDMA-Based Cellular Networks

Abdelbaset S. Hamza, *Student Member, IEEE,* Shady S. Khalifa, *Student Member, IEEE,*
Haitham S. Hamza, *Member, IEEE,* and Khaled Elsayed, *Senior Member, IEEE*

Abstract—Orthogonal Frequency Division Multiplexing Access (OFDMA) has been increasingly deployed in various emerging and evolving cellular systems to reduce interference and improve overall system performance. However, in these systems Inter-Cell Interference (ICI) still poses a real challenge that limits the system performance, especially for users located at the cell edge. Inter-cell interference coordination (ICIC) has been investigated as an approach to alleviate the impact of interference and improve performance in OFDMA-based systems. A common ICIC technique is *interference avoidance* in which the allocation of the various system resources (e.g., time, frequency, and power) to users is controlled to ensure that the ICI remains within acceptable limits. This paper surveys the various ICIC avoidance schemes in the downlink of OFDMA-based cellular networks. In particular, the paper introduces new parameterized classifications and makes use of these classifications to categorize and review various static (frequency reuse-based) and dynamic (cell coordination-based) ICIC schemes.

Index Terms—OFDMA, Long Term Evolution (LTE), Inter-cell Interference coordination (ICIC), Frequency Reuse.

I. INTRODUCTION

NEXT generation cellular systems promise significantly higher cell throughput and improved spectral efficiency as compared to existing systems such as GSM, EDGE, and HSPA+ (High-Speed Packet Access Release 7). For example, system performance requirements for the 3rd Generation Partnership Project (3GPP) Long Term Evolution (LTE) of UMTS [1] and LTE-Advanced [2] target significant improvements in cell-edge spectral efficiency and peak transmission rates that can reach, respectively, 0.04-0.06 bps/Hz/cell and 100 Mbps and beyond. In order to achieve these targets, dense frequency reuse of the scarce radio spectrum allocated to the system is needed. Efficient use of radio spectrum is also important from a cost-of-service point of view, where the number of served users is an important factor. However, as the frequency reuse increases, so does the interference caused by other users using the same channels. Therefore, interference becomes a decisive factor that limits the system capacity, and hence, the suppression of such interference becomes of a particular importance to the design of next generation cellular networks.

Generally speaking, cellular mobile communication systems suffer from two major classes of interference, namely, *intra-cell* interference and *inter-cell* interference. In the former, interference is caused between frequency channels within the same cell due to adjacency of both frequencies and power leakage from one channel to an adjacent channel. In the latter, however, interference is caused between a frequency channel in one cell and the same frequency channel used in another adjacent cell.

In the downlink of the emerging cellular systems such as Worldwide Interoperability for Microwave Access (WiMAX), LTE and LTE Advanced, Orthogonal Frequency Division Multiplexing (OFDM) or Orthogonal Frequency Division Multiple Access (OFDMA) was selected to reduce interference and to efficiently meet their high performance requirements [50].

In OFDMA systems the system spectrum is divided into a number of channels; each channel consists of a cluster of a number of consecutive orthogonal OFDM subcarriers (See Figure 1) [50]. As subcarriers are orthogonal, intra-cell interference is highly reduced. Time is divided into slots consisting of a few consecutive OFDM symbols. The number of OFDM symbols per slot is defined according to the cyclic prefix type used to mitigate inter-symbol interference. In the case of the normal cyclic prefix, the number of symbols per slot is 6, otherwise it is 7. The smallest resource unit that can be scheduled is the *resource block* (RB). A resource block is a single channel for the duration of one time slot. A super-frame is constructed by a number of consecutive frames and a frame is constructed by a number of consecutive slots. Depending on the application, one or more RBs can be allocated to a single user at a time. Each RB is assigned exclusively to one user at any time within a given cell; however, neighboring cells may reuse the same RB for different users.

In OFDMA systems, the transmission rate of a channel is variable and differs based on the user allocated to this channel due to the use of Adaptive Modulation and Coding (AMC). Each enhanced NodeB (eNB) collects the Channel Quality Indicator (CQI) reports fed back from the users which are derived from the downlink received reference signal quality. The CQI is then used to determine the Modulation and Coding Scheme (MCS) for a channel. The modulation schemes ranges from the robust low rate QPSK scheme to the high rate but more error prone 64-QAM scheme. Same MCS is used for all sub-carriers in a RB allocated for a given user though different MCS can be allocated to different resource blocks [50]. The channel throughput is determined based on the used
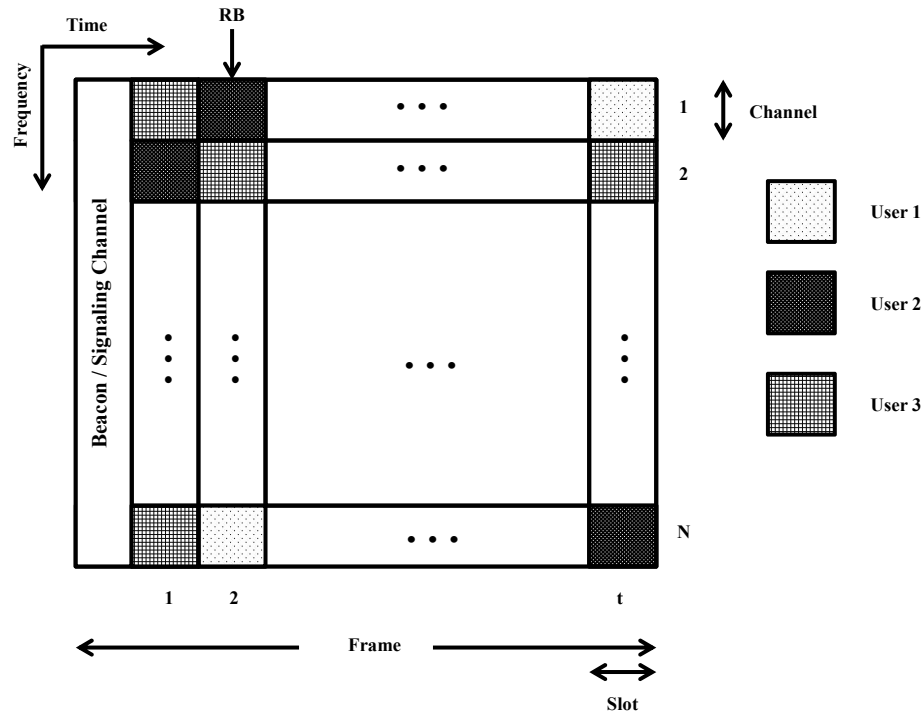
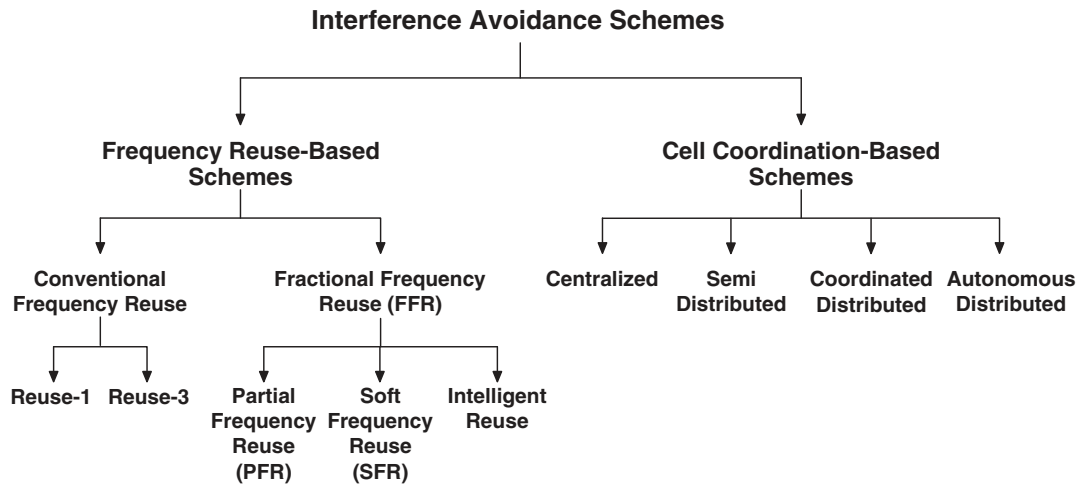Fig. 1. Basic structure of frames in OFDMA Systems.

Fig. 2. Inter-Cell Interference Avoidance Schemes.

MCS (selected based on the channel CQI reported from the user) which is mapped to the Transport Block Size (TBS) that can be used by using the mapping tables in [62]. Since different users perceive different channel qualities, a "bad" channel (due to deep fading and narrowband interference) for one user may still be favorable to other users. Thus, OFDMA exploits the multi-user diversity by avoiding assigning "bad" channels, which is an important feature in OFDMA [30].

Even with almost no intra-cell interference, inter-cell interference (ICI) still presents a great challenge that greatly limits the system performance, especially for users located at the cell edge. In OFDMA systems, ICI is caused by the collision between resource blocks [49]. With such collision model, the overall system performance is determined by the collision probabilities and the impact of a given collision on the Signal to Interference and Noise Ratio (SINR) associated with the colliding resource blocks. Inter-cell interference coordination (ICIC) mechanisms aim at reducing the collision probabilities and at mitigating the SINR degradation that such collisions may cause in order to improve the system performance and increase the overall bit rates of the cell and its cell edge users. Generally speaking, ICIC techniques can be classified into *mitigation* and *avoidance* techniques.

In interference *mitigation*, techniques are employed to reduce the impact of interference during the transmission or after the reception of the signal. In the literature [3]- [7], a wide range of techniques is presented in order to improve the throughput of the cell-edge users by reducing or suppressing

the ICI. Interference mitigation techniques include [6]: (1) Interference randomization, where some cell-specific scrambling, interleaving, or frequency-hopping (spread spectrum) [6], (2) interference cancelation: where the interference signals are detected and subtracted from the desired received signal, or if multiple antenna system is employed, the receiver can select the best quality signal among the various received signals [4], (3) adaptive beamforming: where the antenna can dynamically change its radiation pattern depending on the interference levels.

Interference *avoidance* schemes represent the frequency reuse planning algorithms used by the network elements to restrict or allocate certain resources (in both frequency and time domains) and power levels among users in different cells. The objective of these frequency reuse planning algorithms is to increase the SINR, and hence, allow the system to support as many users as possible. These frequency reuse planning algorithms must satisfy the power constraint in each cell by ensuring that the allocated transmission power of an eNB does not exceed the maximum allowable power. A fundamental concept common to most interference avoidance schemes is to classify users in the cell based on their average SINR to a number of users' classes (also known as "cell regions"). Interference avoidance schemes then apply different reuse factors to the frequency band used by the different classes of users (i.e, to different cell regions).

Various avoidance (allocation) techniques have been studied in the literature under various traffic conditions and network structures. Schemes under this category can be classified along several orthogonal dimensions to mainly differentiate between *static* versus *dynamic*, and *centralized* versus *distributed* techniques. Moreover, avoidance schemes differ with respect to the resources that are being allocated/ coordinated between users, and whether various power levels need to be used at different locations in the cell. Figure 2 depicts the various types of interference avoidance schemes.

In *Static allocation schemes* the resources allocated to each cell and users class are computed and evaluated during the radio planning process and only long-term readjustments are performed during the operation of the network. Thus, the set of sub-carriers and the power levels allocated to each cell and user class is fixed (i.e static). *Static allocation schemes* are relatively easy to implement as they require no frequent interaction among involved base stations. However, since it is based on static frequency reuse, once this allocation scheme is used, it is not easy to perform modifications to the major frequency distributions [3]. Consequently, this scheme is not adaptive to meet dynamic demand changes per sector as it adapts to the cell loads only by changing power used over different sub-carriers. To confront this limitation, *cell coordination based schemes* were investigated, where coordination among neighboring base stations, on both sub-carriers and power levels is used. As a result, such schemes allow for efficient adaptation to the variations in cell loads.

A. Motivation and Scope

Motivated by the pressing need for developing high-speed high-performance cellular networks, the last few years have witnessed an increasing interest in the research community to develop various inter-cell interference avoidance schemes for OFDMA-based cellular networks. As a result, several research papers have been published and, to the best of our knowledge, there exists no comprehensive survey that investigates such a wide range of ICIC avoidance schemes. Moreover, due to the large number of published work in this area, there have been several confusions and contradictions between the various schemes either in their naming conventions or their operational principles. The following are some examples that can create such confusions:

- Some published work uses the notion of "Partial Frequency Reuse (PFR)" [12] while others use "Fractional Frequency Reuse with full isolation (FFR-FI)" [8] to refer to the same scheme.
- Some published work refers to the well known "Reuse-3" scheme as "Hard frequency reuse" [57].
- The notion of "Soft Frequency Reuse (SFR)" was originally proposed in [9] with a particular definition, whereas in [11] a different scheme was introduced with the same name of "Software Frequency Reuse (SFR)".

In addition to the above, there has been a large number of approaches proposed for dynamic interference avoidance (e.g., power aware, fixed power, coordinated, distributed, autonomous, eNB initiated, UE initiated,etc.). These techniques are difficult to comprehend and compare as they focus on the ICI problem from different perspectives.

Based on the above, we believe that there is a need for a survey to collate and present, in a systematic way, current advances in the area of ICIC. Moreover, we believe that such a survey should attempt to resolve the confusion and ambiguity by providing a more precise classification criterion that does not depend merely on the traditional naming convention that has been used so far in the research community. Accordingly, this paper surveys various techniques proposed for static (frequency reuse based) as well as dynamic (coordination-based) inter-cell interference avoidance schemes for the downlink of OFDMA-based cellular networks. In addition, the paper proposes novel parameterized classification approaches to express various static and dynamic avoidance schemes based on their structure and operation in order to reduce ambiguity and increase understandability.

It is worth noting that, even though recent collaborative communications technologies, such as multi input multi output (MIMO), coordinated multi-point (CoMP), and Relay transmission schemes, can contribute to the solution of interference problem; however, these areas have wider scope and may require a separate focused survey, and hence, they are not included in the scope of this work.

B. Related Work

This section briefly reviews main survey papers related to interference avoidance schemes.

In [57], Zhang proposed a classification for interference avoidance schemes with four categorizes. A scheme is allocated to one of these four categories based on its degrees of freedom to adapt to network conditions. The proposed four categories are:

- *Static Schemes*: where in design time, the best values for the different parameters (power ratio allocated to each user class, number of sub-bands allocated to each user class, frequency allocated to each cell) are determined based on full traffic load scenarios and then these values are kept fixed.
- *Low level dynamic Schemes*: As the best values for the different parameters may not always be "best" with different traffic loads, *Low level dynamic Schemes* uses several pre-planned sets of best values for the different traffic loads and varied distributions of users. Given that base stations (BSs) can know the total number of user and there are reliable and efficient connections between BSs, a scheme can switch based on the traffic load between two or more sets of best values each optimized for a certain traffic load.
- *Intermediate level dynamic schemes*: Given the serving-user's quantity in each cell and locations of users in its own cell data available to the BSs, BSs calculates the best values for the different parameters to escape the limitation of using one of the pre-planned best value sets in *Low level dynamic Schemes*.
- *High level dynamic schemes*: Unlike the above categorizes that depend only on the user's quantity, schemes in this category require the availability of the channel condition information. *High level dynamic schemes* works similarly to *Intermediate level dynamic schemes* to calculate the best values for power ratio, the sub band number and allocation of frequency but it also calculates the number of sub channel to be allocated to each user based on its channel condition. However, there were no implementation or evaluation to *High level dynamic schemes* in the paper.

In his paper [57], Zhang only introduced the static interference avoidance schemes, provided no analysis of their performance and didn't use the proposed categorizes to classify any of the published work. Regarding the proposed schemes classification, the analysis showed that as the degrees of freedom increases the total throughput and 10% throughput increase. However, while *low level dynamic schemes* and *Intermediate level dynamic schemes* can provide better performance than the *Static schemes*, they are not justified as nowadays users can send channel condition reports to the base stations on regular relatively small intervals which makes *High level dynamic schemes* much more logical than both of them and so there would be no need for either the *low level dynamic schemes* or *Intermediate level dynamic schemes*.

Even though, multi-cell interference avoidance in OFDMA systems has been for a couple of year now a hot research area with a large number of recently published work, as far as we know, there are no published surveys covering the multi-cell interference avoidance schemes in OFDMA systems. The only comprehensive survey that is related to this research area is the work published by Katzela *et al.* in 1996 [48], which surveys the various channel allocation schemes and a number of channel reuse schemes with limited focus on the ICI problem.

Katzela *et al.* in [48] classify the channel allocation schemes into the following three categories:
- *Fixed Channel Allocation (FCA)*: In FCA, a set of nominal channels is permanently allocated to each cell for its

exclusive use. Where Channels can be allocated to cells either *uniformly* (equal shares) or *nonuniformly* (based on expected traffic loads) with the option of allowing cells to borrow channels from one another.
- *Dynamic Channel Allocation (DCA)*: In DCA, all channels are kept in a central pool and are assigned dynamically to cells when requested and then returned back to the central pool when became idle. The main idea of DCA schemes is to allocate a channel that minimizes the system cost provided that certain interference constraints are satisfied. Based on information used for channel assignment, DCA schemes can be classified either as *call-by-call* (use only current channel usage conditions) or *adaptive* (use previous as well as current channel usage conditions). Based on the type of control employed, DCA schemes can be classified either as *Centralized* (a centralized controller assigns channels to users) or *Distributed* (base stations assigns channels to users). *Distributed* DCA schemes can be either *cell-based* (base stations use local information collected from users and the exchanged information from other base stations) or *adaptive* (base stations rely only on the signal strength measurements collected locally from its users).
- *Hybrid Channel Allocation (HCA)*: HCA presents a mixture between FCA and DCA where the total number of channels available is divided into *fixed* and *dynamic* sets. The *fixed* set is assigned as in the FCA schemes while the *dynamic* set is shared by all cells.

The schemes covered in [48] were evaluated on multi-cell traditional cellular networks but not on OFDMA systems which make the schemes doubtful to perform as stated in the survey when imported to any of the emerging OFDMA systems such as WiMAX, LTE or LTE-Adv [30]. This is due to several reasons. Firstly, unlike traditional cellular networks that assumes a predetermined SINR threshold (for homogeneous applications such as voice), modern data networks utilize adaptive modulation which makes channel assignment decision non-binary from SINR standpoint. UEs employ different modulation and coding schemes with different SINR, thus different throughputs (or achievable rates) are obtained at different SINR levels. Secondly, UEs are frequency selective and their data rate requirements are also different. Finally, the emerging OFDMA systems have put aggressive performance targets that were not planned for traditional cellular networks to handle. Emerging OFDMA systems have triggered a new wave of studies both within the academia and the industry for radio resource management in general and interference coordination in particular that were not included in [48].

While our focus is on the downlink, Yaacoub *et al.* present in [18] a survey of resource allocation and scheduling schemes for the *uplink* channels in OFDMA wireless networks. As the main concern of this survey was resource scheduling, single cell was considered most of the time while ICI in multi-cell scenarios was not heavily discussed.

In [42] Sadr *et al.* presented a survey on resource allocation algorithms for the downlink of multi-user ODFM system. However in this survey a single cell was assumed, thus inter-cell interference and inter-cell interference coordination for the downlink were not discussed.

C. Notations

Because various OFDMA-based cellular technologies make use of different terminologies and definitions for various components in the network, we unify these terminologies and use them throughout the paper to avoid confusion and improve clarity.

The term `User Equipment` (UE) is used to refer to the network end users. We use the term `eNB` to refer to the network element used by the UEs to access the network. An eNB can be a Base Station (BS) or an Access Point (AP). For the central entity controlling a number of eNBs, the term `Radio Network Controller` (RNC) is used. The term `channel` is used to indicate a resource unit to be assigned or restricted to users. A channel can be a `resource block` (RB), a `sub-channel`, `sub-band`, `chunk`, or a `sub-carriers group`. These terms will be used throughout the paper interchangeably.

The term `cell` and `sector` will be used interchangeably throughout the rest of the paper. Finally, The term `inter-cell interference coordination` `(ICIC) schemes` will be used throughout the rest of the paper to refer to the interference avoidance schemes.

D. Paper Organization

The remainder of this paper is organized as follows. In Section II, a classification of frequency reuse-based schemes is presented and various schemes are explained. Section III presents a classification for various coordination-based interference avoidance schemes and explains some of these schemes. A discussion on future research directions is presented in Section IV. Finally a summary is given in Section V.

II. STATIC ICIC: FREQUENCY REUSE-BASED SCHEMES

One of the fundamental techniques to deal with the ICI problem is to control the use of frequencies over the various channels in the network. Frequency reuse-based schemes include: conventional frequency planning schemes (Reuse-1 and Reuse-3), fractional frequency reuse (FFR), partial frequency reuse (PFR), and soft frequency reuse (SFR).

Despite their differences, all frequency reuse-based schemes need to specify: (1) the set of channels (sub-bands) that will be used in each sector/cell, (2) the power at which each channel is operating, and (3) the region of the sector/cell in which this set of channels are used (e.g., cell-centre or cell-edge). Different schemes define different values and approaches for these various parameters.

Accordingly, we can identify a unified structured description for any frequency reuse-based scheme. We believe that such a structured description will not only simplify the expression of various schemes, but it will also reduce ambiguity in understanding some of the subtle schemes reported in the literature. To this end, in the following subsections, we introduce a new classification model, and use this model to explain some of the key frequency reuse-based schemes reported in the literature.
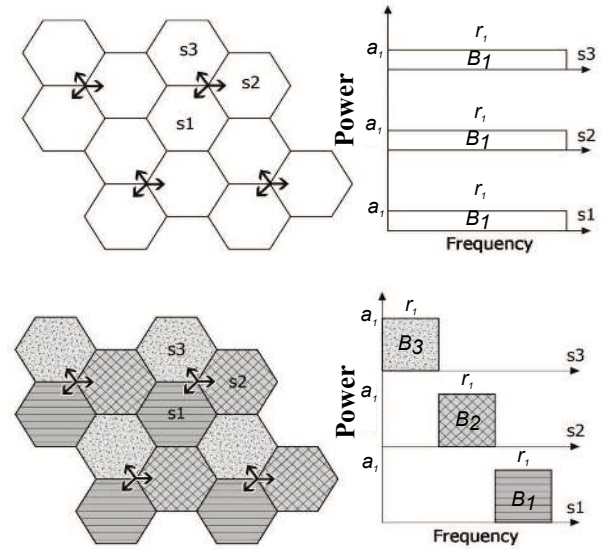

Fig. 3. Conventional Frequency Planning (a) Reuse-1 (b) Reuse-3.

A. Definitions and Notation

The following are the basic parameters that we use to develop the proposed classification model.

- $\mathcal{S} = \{s_1, s_2, s_3\}$: the set of sectors per cell
- $\mathcal{B} = \{B_1, B_2, ..., B_k\}$: the set of consecutive frequency bands that constitute the frequency spectrum in each cell such that $B_i = (b_{i-1}, b_i]$, where b_i refers to frequency i.
- $\mathcal{R} = \{r_1, r_2, ..., r_m\}$ the set of co-centric rings that constitute the cell such that: $\forall r_j \in R, j > 1, r_j$ represents a ring that is bounded between the two radii ρ_{j-1}, ρ_j. A ring r_j defines a user class (center, edge, ...) that can be allocated to a set of specified channels. While ρ_j defines the threshold after which a user is not classified as r_j. For $j = 1, r_j$ represents the cell center users.
- $\alpha = \{a_1, ..., a_n\}$: the set of ascending power levels used within the sub-bands of a cell with respect to the maximum available power in the system. That is, $a_z \in \alpha = \frac{P_z}{P_{max}}$, where P_z is the power level used in a particular sub-band, and P_{max} is the maximum power used in the system. A sub-band that is not used in a particular cell will be assigned a power level 0.

A particular frequency reuse-based scheme can be described by defining the cardinality and elements of the above parameters, then, each sector/cell, $s_i \in \mathcal{S}$, can be expressed as follows:

$$s_i = \{B_k(a_n, r_m) : 1 \le k \le |\mathcal{B}|; a_n \in \alpha; r \in \mathcal{R}\}$$

B. Conventional Frequency Planning

The simplest scheme to allocate frequencies in a cellular network is to use a frequency reuse factor (FRF) of 1, that is, all available frequency spectrum is reused in each sector without imposing any restrictions on frequency resource usage or power allocation [Figure 3-(a)]. The *reuse-1 scheme* can be described as follows: $\mathcal{B} = \{B_1\}$, $\mathcal{R} = \{r_1\}$, and $\alpha = \{a_1\}$. The three sectors are identical, hence:

$$s_1 = s_2 = s_3 = \{B_1(a_1, r_1)\}.$$

Apparently, this scheme allows for achieving the high peak data rate. However, this comes at the cost of suffering the worst case inter-cell interference levels, especially for cell edge users. This in turn, will greatly limit the performance of these users, leading to an overall lower spectral efficiency.

To reduce the ICI level resulted in the reuse-1 scheme above, the whole frequency band can be divided into three equal but orthogonal sub-bands. Adjacent sectors will be allocated different sub-bands [Figure 3-(b)]. This scheme is known as *reuse-3*, and it can be described as follows: $\mathcal{B} = \{B_1, B_2, B_3\}$, $\mathcal{R} = \{r_1\}$, $\alpha = \{a_1\}$. Accordingly, the three sectors of this scheme can be expressed as follows:

$$s_1 = \{B_1(0,0), B_2(0,0), B_3(a_1, r_1)\}$$
$$s_2 = \{B_1(0,0), B_2(a_1, r_1), B_3(0,0)\}$$
$$s_3 = \{B_1(a_1, r_1), B_2(0,0), B_3(0,0)\}$$

This clustering obviously leads to an improved (lower) inter-cell interference; however, this comes at the cost of very low bandwidth utilization due to the restrictions imposed on the reuse of the available resources. In fact, only one third of the resources are utilized in each sector.

It appears from the above discussion that conventional frequency planning schemes represent the lower and upper bounds on the interference as well as resource utilization in the network. While reuse 1 does not employ any interference coordination, reuse 3 can be regarded as an extreme case of partition based static interference coordination.

C. Fractional Frequency Reuse (FFR)

To avoid the shortcomings of the conventional frequency reuse schemes, the fractional frequency reuse (FFR) scheme is introduced to achieve a FRF between 1 and 3. FFR divides the whole available resources into two subsets or groups, namely, the *major* group and the *minor* group. The former is used to serve the cell-edge users, while the latter is used to cover the cell-center users. Generally speaking, the FFR scheme can be divided into three main classes:

1) *Partial Frequency Reuse (PFR) Schemes:* in these schemes a common frequency band is used in all sectors (i.e., with a frequency reuse-1) with equal power, while the power allocation of the remaining sub-bands is coordinated among the neighboring cells in order to create one sub-band with a low inter-cell interference level in each sector.

2) *Soft Frequency Reuse (SFR) Schemes:* in these schemes, each sector transmits in the whole frequency band. However, the sector uses full power in some frequency sub-bands while reduced power is used in the rest of the frequency band.

3) *Intelligent Reuse Schemes:* in these schemes, band allocated to different sectors expands and dilates based on the existing workloads. These schemes start with a reuse-3 like configuration at low workloads which can be changed with the increase of workloads to become PFR, SFR or even reuse-1.

In [10] a study that attempts to find an optimum FFR is presented where the problem is formulated as sum-power minimization problem subject to minimum rate constraints in both the regions. The study considers the optimal FFR factor for the cell-edge region, bandwidth assigned to each region
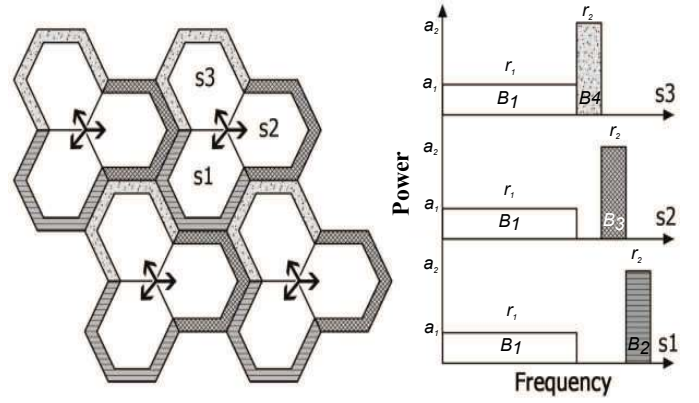

Fig. 4. Fractional Frequency Reuse with Full Isolation (FFR-FI).

and subcarrier and power allocation to all the users in the cell. The key result is that for the same minimum demanded rate for all users, it is found that the power consumed is minimal when the reuse factor used for the cell-edge region is 3. In the following, a detailed discussion of different frequency reuse schemes is presented.

1) Partial Frequency Reuse (PFR): From the above discussion, it is clear that using the same FRF value for the entire cell is not bandwidth-efficient [8]. One way to improve the cell-edge SINR, while maintaining a good spectral efficiency, is to use an FRF greater than unity for the cell-edge regions and an FRF of unity for the cell-center regions [11]. In a homogeneous network, the cell centre regions have equal areas.

The idea of the partial frequency reuse (PFR) is to restrict portion of the resources so that some frequencies are not used in some sectors at all. The effective reuse factor of this scheme depends on the fraction of unused frequency [12].

The PFR is also known as *FFR with full isolation* (FFR-FI), as users at cell-edge are fully protected (isolated) from adjacent cells' interference [8]. An example for sites with 3 sectors is shown in Figure 4. The effective reuse of PFR is greater than one. To see this, consider a system with available bandwidth equal to β. This bandwidth is divided into inner and outer zones with bandwidth equal to β_i and β_0, respectively. Band β_i is used with a reuse factor of 1, and for the tri-sector BSs, the reuse factor for β_0 is usually 3 in the outer zone. In this case, the effective frequency reuse factor is given by $\beta/(\beta_i + (\beta_0/3))$. Therefore, the effective reuse of PFR scheme is always greater than 1 [12].

This scheme can be described as follows: $\mathcal{B} = \{B_1, B_2, B_3, B_4\}$, $\mathcal{R} = \{r_1, r_2\}$, $\alpha = \{a_1, a_2\}$. Accordingly, the three sectors of this scheme can be expressed as follows:

$$s_1 = \{B_1(a_1, r_1), B_2(0,0), B_3(0,0), B_4(a_2, r_2)\}$$
$$s_2 = \{B_1(a_1, r_1), B_2(0,0), B_3(a_2, r_2), B_4(0,0)\}$$
$$s_3 = \{B_1(a_1, r_1), B_2(a_2, r_2), B_3(0,0), B_4(0,0)\}$$

A numerical method for calculation of interference generated by co-channel cells is proposed and discussed in [13] [14]. The level of co-channel interference in three different scenarios is compared, in particular, cellular system with universal frequency reuse, cellular system with reuse-3 and cellular system with implemented ICIC based on fractional frequency reuse. Analysis shows that the interference experi-
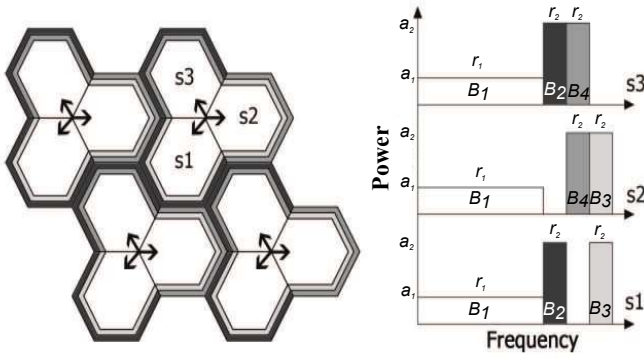
Fig. 5. PFR with only one interference in the worst case.

Fig. 6. Frequency Reuse: (a) Reuse-1 in the cell-centre, (b) Partial frequency reuse in the cell-centre.

enced by users in their own cells is almost two times smaller when using fractional frequency reuse instead of frequency reuse factor 3 and approximately three times smaller than universal frequency reuse case.

A novel fractional frequency reuse scheme combined with interference suppression for orthogonal frequency division multiple access (OFDMA) networks is introduced in [20]. The *PFR with only one interference in the worst case scheme* (Figure 5) ensures maximum of one-type interferer, that is only users in the neighboring cells using the same band will cause interference to the cell edge users, and hence, it was possible to suppress this interference by using interference exploitation techniques. Results indicate a reduction in power at no cost of increased complexity. This FFR scheme can be classified as a variation of the PFR scheme, and it can be described as follows: $\mathcal{B} = \{B_1, B_2, B_3, B_4\}$, $\mathcal{R} = \{r_1, r_2\}$, $\alpha = \{a_1, a_2\}$. Accordingly, the three sectors of this scheme can be expressed as follows:

$$s_1 = \{B_1(a_1, r_1), B_2(a_2, r_2), B_3(0, 0), B_4(a_2, r_2)\}$$
$$s_2 = \{B_1(a_1, r_1), B_2(0, 0), B_3(a_2, r_2), B_4(a_2, r_2)\}$$
$$s_3 = \{B_1(a_1, r_1), B_2(a_2, r_2), B_3(a_2, r_2), B_4(0, 0)\}$$

2) Soft Frequency Reuse (SFR): The PFR scheme may result in under-utilization of available frequency resources due to its strict no-sharing policy. Soft Frequency Reuse (SFR) was proposed in [5] [9] to present a balance between the FRF and the PFR schemes. It avoids the high ICI levels associated with the unity FRF configurations, while providing more flexibility to the PFR scheme. The term soft reuse is due to the fact that effective reuse of the scheme can be adjusted by the division of powers between the frequencies used in the centre and edge bands.

SFR makes use of the concept of zone-based reuse factors in the cell-center and cell-edge areas. Unlike the PFR; however, frequency and power used in these zones are restricted. In particular, a frequency reuse factor of 1 is employed in the central region of a cell, while frequency reuse factor greater than 1 at the outer region of the cell close to the cell edge.

For example, consider the 3-sector cell sites shown in Figure 6, the cell-edge band (major band) uses 1/3 of the available spectrum which is orthogonal to those in the neighboring cells and forms a structure of cluster size of 3. The cell-centre band (minor band) in any sector is composed of the frequencies used in the outer zone of neighboring sectors. According to
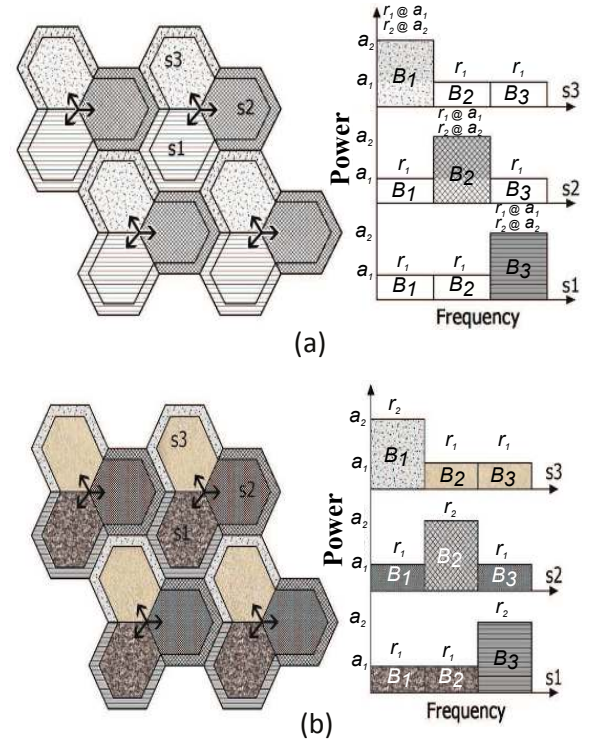
the original contribution in which the SFR is proposed [9], the major band can be used in the cell-centre as well if it is not occupied by the cell-edge UEs, resulting in a frequency reuse factor of 1 for the inner part of the cell, but the minor band is available to the centre area only [Figure 6-a)]. Each group is assigned transmission power depending on the desired effective reuse factor, such that the major band group is higher than the power of the minor group while keeping the total transmission power fixed. Higher transmit power is used on the major band as shown in the right side of Figure 6.

This scheme can be described as follows: $\mathcal{B} = \{B_1, B_2, B_3\}$, $\mathcal{R} = \{r_1, r_2\}$, $\alpha = \{a_1, a_2\}$. Accordingly, the three sectors of this scheme can be expressed as follows:

$$s_1 = \{B_1(a_1, r_1), B_2(a_1, r_1), B_3(a_1, r_1) \cup B_3(a_2, r_2)\}$$
$$s_2 = \{B_1(a_1, r_1), B_2(a_1, r_1) \cup B_2(a_2, r_2), B_3(a_1, r_1)\}$$
$$s_3 = \{B_1(a_1, r_1) \cup B_1(a_2, r_2), B_2(a_1, r_1), B_3(a_1, r_1)\}$$

However, the authors in [11] define the SFR differently. In their work, they refer to the SFR as a scheme in which the available bandwidth is divided into orthogonal segments, and each neighboring cell is assigned a cell-edge band, where a higher power is allowed on the selected cell-edge band, while the cell-centre UEs can still have access to the cell-edge bands selected by the neighboring cells, but at a reduced power level. In this way, each cell can utilize the entire bandwidth while reducing the interference to the neighbors [Figure 6-b)]. A less ICI at the cell-edge is achieved at the expense of spectrum utilization.

This scheme can be described as follows: $\mathcal{B} = \{B_1, B_2, B_3\}$, $\mathcal{R} = \{r_1\}$, $\alpha = \{a_1, a_2\}$. Accordingly, the three sectors of this scheme can be expressed as follows:

$$s_1 = \{B_1(a_1, r_1), B_2(a_1, r_1), B_3(a_2, r_2)\}$$

$$s_2 = \{B_1(a_1, r_1), B_2(a_2, r_2), B_3(a_1, r_1)\}$$
$$s_3 = \{B_1(a_2, r_2), B_2(a_1, r_1), B_3(a_1, r_1)\}$$

In [5], a parameter called "Power Ratio" is defined as the ratio between transmit power limitation of minor sub-carriers and major sub-carriers. Adjusting the power ratio from 0 to 1 effectively moves the reuse factor from 3 to 1. Therefore, SFR is seen as a compromise between reuse 1 and 3 in a network with tri-sector BSs. UTs are categorized into cell-edge and cell-centre based on user geometry determined by the received signal power (averaged over multipath fading) taking into account the large-scale path-loss, shadowing, and antenna gains [12].

Simulation results reported in [5] [9] show that: if the power ratio equals one (i.e., major and minor subcarriers are given same power), the cell-edge bit rate equals one third of the cell-edge bit rate in case of the universal reuse factor (i.e., reuse-1). As the power factor decreases towards 0, the total cell throughput decreases as well. Also, the throughput of the inner zone decreases as well. However, the cell-edge throughput increases due to the increasing transmission power for cell-edge users and the mitigation of co-channel interference. The above discussion can lead to the general conclusion that the SFR scheme can improve the SINR of the cell-edge UEs using a greater than unity FRF, while degrading the SINR of the cell-centre UEs. This degradation is due to the overlap in frequency resources between the cell-edge band of the neighboring cells, and the cell-centre band of the serving cell. However, as the ICI is not as dominant and important for the cell-centre UEs as for the cell-edge UEs, and since a cell-center UEs SINR is typically much higher than unity, the cell-centre UEs spectral efficiency increases only logarithmically with SINR. However, for the cell-edge UEs with SINR value much less than unity, the spectral efficiency increases almost linearly with SINR. This leads to a cell-edge performance improvement almost linear with SINR while the degradation to the cell-centre UEs is logarithmic with SINR. In SFR, the power ratio between the cell-edge band and the cell-centre band can be an operator-defined parameter, thereby increasing the flexibility in system tuning.

In [15], the performance of the SFR with partial frequency reuse at the cell centre for large scale networks in realistic radio environments and with irregular cell patterns is investigated. According to simulations, two key conclusions are drawn. SFR's parameters have to be carefully selected and optimized since any improvement for the cell edge users comes at the expense of performance of the cell centre users. Therefore, it was suggested that the SFR is better used for resolving interference issues at some specific areas, at which the performance reduction at the center zones is much significant compared to the improvement in the cell edge, rather than being used in the entire network. It is also found that the cell-edge performance is sensitive to the bandwidth allocated; an interesting result is that the two sub-bands have better performance over the conventional three sub-bands. A recommendation for further work is then presented for performing comparison studies between the SFR scheme and other schemes. Also, the usage sub-bands of unequal sizes is recommended. This can better adapt the reuse pattern to the cell layout.
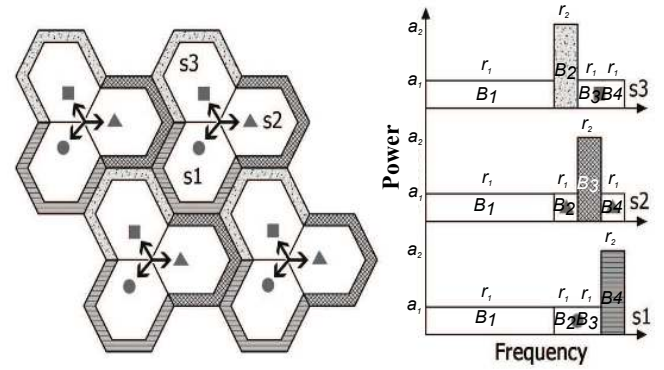

Fig. 7. Soft Fractional Frequency Reuse (SFFR).

Soft Fractional Frequency Reuse (SFFR) is an enhanced SFR scheme. The PFR and SFR schemes can improve the throughput for the cell-edge users by reducing the ICI experienced by users in that region. However, both schemes may lead to a lower cell throughput as compared to the conventional reuse-1 scheme. The PFR scheme does not utilize the whole available frequency bandwidth, and thus, it has a lower cell throughput as compared to reuse-1 scheme. Moreover, although SFR can make use of the overall frequency band available in the cell, and thus, increase the overall system capacity compared to that of the PFR; however, the overall system capacity of SFR maybe lower than that of reuse one scheme.

Soft FFR (SFFR) scheme has been proposed as a way to improve the overall cell throughput of FFR [16]. Unlike the PFR that does not make use of the sub-bands allocated to the outer region in the adjacent cells, the Soft FFR scheme utilizes these sub-bands for the inner UEs, but with low power levels (See Figure 7). As a result, the SFFR is similar to the SFR in that both adopt a non-uniform power profile (it uses high power levels for some sub-bands and low power levels for others). Unlike the SFR; however, the Soft FFR uses the common sub-band which can enhance the throughput of the inner users. This scheme can be described as follows: $\mathcal{B} = \{B_1, B_2, B_3, B_4\}$, $\mathcal{R} = \{r_1, r_2\}$, $\alpha = \{a_1, a_2\}$. Accordingly, the three sectors of this scheme can be expressed as follows:

$$s_1 = \{B_1(a_1, r_1), B_2(a_1, r_1), B_3(a_1, r_1), B_4(a_2, r_2)\}$$
$$s_2 = \{B_1(a_1, r_1), B_2(a_1, r_1), B_3(a_2, r_2), B_4(a_1, r_1)\}$$
$$s_3 = \{B_1(a_1, r_1), B_2(a_2, r_2), B_3(a_1, r_1), B_4(a_1, r_1)\}$$

In [17], several variations of the power profile used in the Soft FFR scheme are investigated. Several interesting observations were presented and can be summarized as follows:

- Transmission power level of the common part does not have a significant influence on the overall cell throughput. Accordingly, the total transmission power used in the cell can be reduced by minimizing the power level of this common part without impacting the required cell throughput level.
- Transmission power level in the outer region has a direct impact on the throughput of that region. In particular, it was observed that the throughput of this region is directly proportional to its power, and inversely proportional to the inner region's throughput. As a result, according to the throughput requirements in the outer region, the power
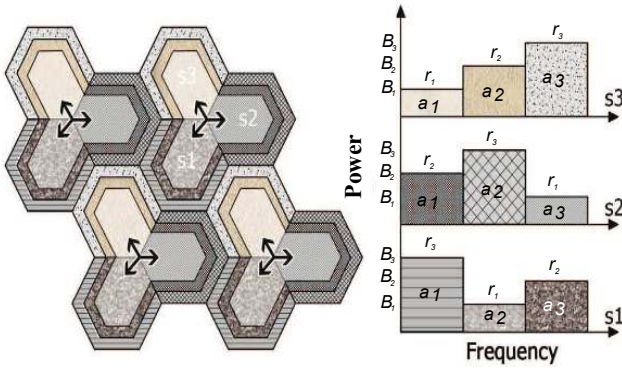
Fig. 8. FFR with multiple user class.

Fig. 9. Low spectrum efficiency problem in SFR.

consumption for the outer region can be reduced, while maintaining high overall system throughput.

Downlinks transmit power allocation in soft FFR under two different coordination cases, namely; *loosely* and *tightly* coordinated cells are studied [19]. In the loosely coordinated cells case, the sub-band transmit powers are allocated so that the cell edge user meets the required throughput. The loss in average cell throughput can be reduced by configuring appropriate number of sub-bands for inner and outer regions. However, in the tightly coordinated cells, sub-band power allocation can be changed packet by packet in each scheduling period. It is found that in this cell coordination case, the loss of spectral efficiency can be minimized regardless of the number of sub-bands due to its fast coordination.

A *FFR with multiple user class scheme* that deals with different user classes is presented in [21] [22]. The scheme uses an approach similar to that of the *PFR with only one interference in the worst case scheme* [20] where cells are divided into a number of concentric zones, each with a different frequency reuse factors as illustrated (Figure 8). However, unlike the *PFR with only one interference in the worst case scheme* where some bands are restricted, under this scheme, the cell uses the entire band but under different power level restrictions based on the type of UEs. Central UEs are served first with the low power sub-band (if this is not enough, next sub-bands can be used, but should still maintain the low power level). Next, intermediate and finally cell edge UEs are served with the same criterion. This scheme can be described as follows: $\mathcal{B} = \{B_1, B_2, B_3\}$, $\mathcal{R} = \{r_1, r_2, r_3\}$, $\alpha = \{a_1, a_2, a_3\}$. Accordingly, the three sectors of this scheme can be expressed as follows:

$$s_1 = \{B_1(a_3, r_3), B_2(a_1, r_1), B_3(a_2, r_2)\}$$
$$s_2 = \{B_1(a_2, r_2), B_2(a_3, r_3), B_3(a_1, r_1)\}$$
$$s_3 = \{B_1(a_1, r_1), B_2(a_2, r_2), B_3(a_3, r_3)\}$$

3) Intelligent Reuse: Incremental Frequency Reuse (IFR): Under the SFR scheme, cell edge users have a maximum of one third of the entire bandwidth to utilize. However; typically, cellular systems have more cell edge users than cell center users. Thus, SFR may result in low spectrum efficiency. Moreover, as shown in Figure 9, under the SFR scheme, co-channel interferences may increase even under low traffic load situation, while there are still sub-channels in idle and underutilized in the system. This 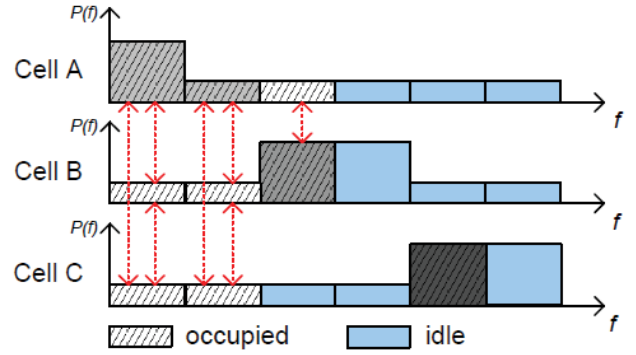is due to the fact that resource allocation of all cells under the SFR scheme starts always from the first sub-channel up. Again, this may reduce the spectrum utilization efficiency.

In addition, results on the usage of SFR showed that the cell throughput is even lower to the conventional reuse-1 scheme when loading factor is over 0.5 [23]. This is because under the SFR scheme, at most one third of the sub-channels can be used to transmit data with higher power while the remaining two third sub-channels work with lower power, which induces an overall throughput loss. Thus, the SFR ameliorates performance of the cell edge users at the expense of degrading the overall cell capacity [24].

In order to overcome some of the shortcomings of the conventional SFR scheme discussed above (low spectrum efficiency, increased co-channel interferences at low loading traffic, and loss of cell capacity system when system is over half-full loaded), in [23], Kim *et al.* proposed the concept of Incremental Frequency Reuse (IFR) scheme. IFR attempts to reduce the ICI effectively under low offered traffic, while maintaining the overall system capacity.

Figure 10 illustrates the basic concept of the IFR scheme in a tri-sector cell system with 3 various types of neighboring cells. The only difference between the IFR and the classical reuse-1 is, from which point of the available bandwidth it starts dispensing resources to the users. In an IFR system the directly adjoining cells assign resources from different sub-channels. Cells of type-A occupy resources from the first sub-channel, whereas cells of type-B from one third of the whole bandwidth, and cells of type-C from two third of the bandwidth. They allocate consecutive sub-channels successively along with traffic load increasing until the entire bandwidth is used up. The ICI generated by directly adjoining cells can be avoided completely at low traffic situation, since frequency reuse of the first tier neighboring cells doesn't occur when loading factor below 0.3, and the whole system operates as in the classical reuse-3 system. Effectively, under the IFR scheme, the system operates with increasing traffic load like moving from a reuse-3 system to a reuse-1 system.

Despite the fact that the IFR scheme can overcome most of the limitations inherited in the SFR scheme; however, the
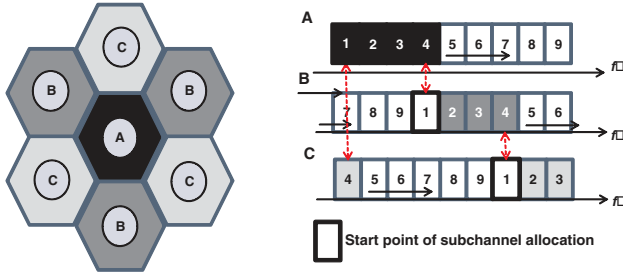
Fig. 10. IFR scheme in a tri-sector cell system.

Fig. 11. Enhanced Fractional Frequency Reuse (EFFR) .

IFR scheme performs better only under low traffic. When the loading factor in the system is above 0.3, the IFR performance is lower than that of the SFR. Simulation results reported in [23]concluded that both the IFR and the SFR schemes do not perform better than the classical reuse-1 scheme in over-middle-load or full-load situations. For IFR, all bandwidth becomes available to all cells resulting in an overall cell capacity equals to that of reuse-1. The SFR scheme performs worse than reuse-1 as two-thirds of the users are allocated to the secondary-band achieving relatively lower throughputs due to the limited transmit power, while only one-third achieving higher throughput. Accordingly, it is concluded that the system capacity cannot be substantively improved by the IFR and the SFR schemes. This scheme can be described as follows: $\mathcal{B} = \{B_1, B_2, B_3\}$, $\mathcal{R} = \{r_1\}$, $\alpha = \{a_1\}$. Accordingly, the three sectors of this scheme can be expressed as follows:

$$s_1 = \{B_1(a_1, r_1), B_2(a_1, r_1), B_3(a_1, r_1)\}$$
$$s_2 = \{B_2(a_1, r_1), B_3(a_1, r_1), B_1(a_1, r_1)\}$$
$$s_3 = \{B_3(a_1, r_1), B_1(a_1, r_1), B_2(a_1, r_1)\}$$

Enhanced Fractional Frequency Reuse (EFFR): To further improve the performance of the IFR and the SFR schemes and overcome their limitations, a scheme called Enhanced Fractional Frequency Reuse (EFFR) was proposed in [24]. EFFR attempts to enhance the system capacity especially under overload situations.

Similar to the IFR scheme, the EFFR scheme defines 3 cell-types for directly neighboring cells in a cellular system, and reserves for each cell-type a part of the whole frequency band named Primary Segment, which is shown in the right part of Figure 11 with thick border.

The Primary Segments among different type cells should be orthogonal. The remaining sub-channels excluding the Primary segments constitute the Secondary Segment. The Primary Segment of a cell-type is at the same time a part of the Secondary Segments belonging to the other two cell-types. Each cell can occupy all sub-channels of its Primary Segment at will, whereas only a part of sub-channels in the Secondary Segment can be used by this cell in interference-aware manner.

The Primary Segment of each cell will be further divided into a reuse-3 part and reuse-1 part. The reuse-1 part can be reused by all types of cells, while reuse-3 part can only exclusively be reused by other same type cells. The reuse-3 sub-channels cannot be reused by directly neighboring cells, that attenuates the co-channel interferences among them and therefore it is specified for the vulnerable cell edge users to take priority of using these sub-channels over cell center users.
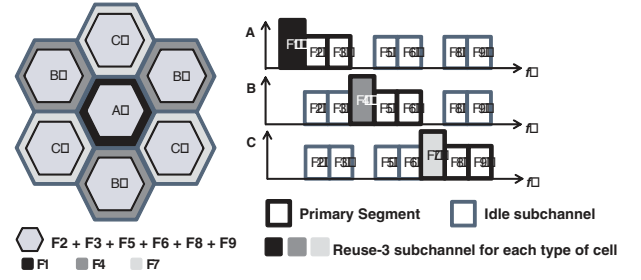
Since a cell acts on the Secondary Segment as a guest, and occupying secondary sub-channels actually reuses the primary sub-channels belonging to the directly neighboring cells, therefore, the Secondary Segment to be reused should be first monitored, then being reused based on the SINR estimation.

Each cell listens on every secondary sub-channel all the time. And before occupation, it makes SINR evaluation according to the gathered channel quality information (CQI) and chooses resources with best estimation value for reuse. If all available secondary resources are either occupied or not good enough to a link, it will give up reusing for this link. This will not lead to resource wasting, which means some resources maybe not reusable for this link, but can be reused by other links. Another gained merit is that it will not generate excessive interference for the neighboring cells which would degrade their performance. So, an upgrade of spectrum efficiency is expected by using the interference-aware-reuse mechanism on the Secondary Segment.

Simulation results for comparing the EFFR scheme with conventional reuse-1, reuse-3, and the IFR schemes show a significant improvement in the overall capacity gains at cell edge as compared to other schemes. This scheme can be described as follows: $\mathcal{B} = \{B_1, B_2, B_3, B_4, B_5, B_6\}$, $\mathcal{R} = \{r_1, r_2\}$, $\alpha = \{a_1, a_2\}$. Accordingly, the three sectors of this scheme can be expressed as follows:

$$s_1 = \{B_1(a_2, r_2), B_2(a_1, r_1), B_3(0, 0), B_4(a_1, r_1), B_5(0, 0),$$
$$B_6(a_1, r_1)\}$$
$$s_2 = \{B_1(0, 0), B_2(a_1, r_1), B_3(a_2, r_2), B_4(a_1, r_1), B_5(0, 0),$$
$$B_6(a_1, r_1)\}$$
$$s_3 = \{B_1(0, 0), B_2(a_1, r_1), B_3(0, 0), B_4(a_1, r_1), B_5(a_2, r_2),$$
$$B_6(a_1, r_1)\}$$

Combined Partial Reuse and Soft Handover: An interesting ICIC scheme that employs both partial frequency reuse (PFR) and soft handover (SH) is proposed in [25]. This proposed scheme differentiates between the cell interior users (CIUs) from the cell edge users (CEUs) using the soft-handover (SH), where a user is considered as a cell edge user if there is at least two cells in its handover list, accordingly the user is allocated one of the sub-channels that are designed to serve this region of the cell. Effectively, this scheme capitalizes on the information already available from the handover algorithm, and thus, it eliminates the complexity of geometry determination based on the duplicate calculation of SINR. In addition, it eliminates the need for extra signaling. Simulation results reported in [25] shows that the combined approach can provide a significant cell edge throughput gain over the conventional

partial frequency reuse scheme. In addition, this scheme is shown to have a low soft handover overhead.

Table I summarizes the various frequency reuse-based schemes using the above classification parameters.

D. Comments on Frequency Reuse-Based ICIC Schemes

In [55] and [56], Gonzalez *et al.* conducted a comparison among a number of static schemes to evaluate the performance of static ICIC schemes in regular cells and realistic irregular cells layouts.

In [55], three static ICIC schemes (PFR [11] Figure 6-(b), FFR with multiple user class [21] Figure 8 and FFR-FI [8] Figure 4) were compared in a regular cells layout. The role of the different elements affecting their spectral efficiency versus fairness tradeoff was investigated. These elements are:

(1) *Threshold to classify users* (ρ_j) which can be either *Class Proportionality* where the SINR thresholds are selected so that each class has the same average number of users or *Bandwidth Proportionality* where the SINR thresholds guarantee that the number of users is proportional to its allocated bandwidth.

(2) *Effect of the power level allocated to different user classes* (α).

(3) *Effect of the number of user groups* ($\mathcal{R}$) which in turn reflect the effect of the size of the bandwidth allocated to each user group ($\mathcal{B}$). With the increase of the number of user groups, the bandwidth allocated to each user group tends to decrease.

(4) *Effect of inter-class interference* which is caused due to using the same band for different user classes at the different cells. Restricting inter-class interference would mean using a reuse-n where $n \geq 3$ for the edge bands as in FFR-FI scheme [8]. Center-users receive ICI of types inter-class (coming from users of different classes (i.e edge users) using the same band in the neighboring cells) and intra-class (coming from users of the same class (i.e center users) using the same band in the neighboring cells). Edge users only receive inter-class interference.

Based on their findings, Gonzalez *et al.* provided the following recommendations:

- The choice of the *threshold to classify users* (ρ_j) has an immediate impact on the scheduler decisions and so on the system performance. When *bandwidth proportionality* is used, the set of cell-edge users becomes smaller leaving more Resource Blocks (RBs) for them to use than the *class proportionality* case leading to a slightly better fairness value. On the other hand, *class proportionality* brings a significant spectral efficiency improvement at the expense of a small fairness degradation.

- Regarding the *effect of the power level allocated to different user classes* (α), it was found that an increase in the power allocated to center-users leads to better values of efficiency no matter which *threshold to classify users* criterion has been selected. The higher the difference in the power assigned to the different user classes, the higher value of fairness and the lower value of efficiency. The reason behind this behavior is that given that the bandwidth assigned to each class is fixed, having a small difference in the power

assigned to the different user classes causes the energy previously assigned to the users having worse channel to move to users with better channel conditions.

- For the *effect of the number of user groups* ($\mathcal{R}$), the analysis showed that as users are grouped into more classes, fairness is improved. Nevertheless, smaller number of user groups increases the spectral efficiency since the allocation of a wider band to the set of users enjoying a better radio channel becomes the predominant effect, especially when the SINR threshold is shifted to higher values.

- The *inter-class interference* can be completely removed but at the expense of a reduction in the available bandwidth at each cell as it can be seen in Figure 4. The advantage of doing so is that higher levels of SINR can be achieved for edge users within the cell as no other cell is using the same band. On the other hand, the reduction in the bandwidth causes a reduction in terms of spectral efficiency. However, the effectiveness in the use of resources is higher compared to schemes with inter-class interference indicated by high values of bits per RB and bits per Watt.

In [56], Gonzalez *et al.* extended their evaluation of static schemes by a comparative study of the performance of four ICIC schemes (reuse-1 Figure 3-(a), reuse-3 Figure 3-(b), SFR [9] Figure 6-(a), FFR-FI [8] Figure 4) in a realistic non-regular cellular layout giving special attention to the efficiency vs. fairness tradeoff. The evaluation scenarios simulated the city of Vienna and its surroundings using the digital elevation model, system layout and propagation data provided by the MORANS initiative [58] which was framed within the European COST 273 Action to provide common system simulation environments so that different researchers can compare results. However, we failed to get such scenario data sets or access the related documentation as they seem have been restricted for public access. Another openly available alternative for real life cellular layouts is the network planning scenarios for Berlin and Lisbon provided by the European Momentum project [59] which contains the city's realistic radio propagation setting. The Momentum project data sets were used in [15] to evaluate the performance of SFR in large networks with irregular cell patterns.

Comparing the scheme's evaluation in realistic irregular cells conducted in [56] to the evaluation conducted on regular shaped cells in [55] led to the following conclusions:

- Regarding the *effect of the power level allocated to different user classes*, the same behavior remains in realistic layouts where an increase in the power allocated to center-users leads to better values of efficiency and lower values of fairness. Also, the higher the difference in the power assigned to the different user classes, the higher value of fairness and the lower value of efficiency.

- The *effect of the number of user groups* in realistic scenarios increases on all performance metrics due to the irregular geometry.

- Also, as users are grouped into less classes, the *effect of the power level allocated to different user classes* increases on all metrics.

In their study, Gonzalez *et al.* concluded that the best performance cannot be obtained by applying traditional ICIC

TABLE I
SUMMARY OF FREQUENCY REUSE-BASED SCHEMES.

Scheme Name(s)	Representation with proposed notation
Reuse-1 Figure 3-(a)	$\mathcal{B} = \{B_1\}$ $\mathcal{R} = \{r_1\}$ $\alpha = \{a_1\}$ $s_1 = s_2 = s_3 = \{B_1(a_1, r_1)\}$
Reuse-3 (Hard frequency reuse [57]) Figure 3-(b)	$\mathcal{B} = \{B_1, B_2, B_3\}$ $\mathcal{R} = \{r_1\}$ $\alpha = \{a_1\}$ $s_1 = \{B_1(0, 0), B_2(0, 0), B_3(a_1, r_1)\}$ $s_2 = \{B_1(0, 0), B_2(a_1, r_1), B_3(0, 0)\}$ $s_3 = \{B_1(a_1, r_1), B_2(0, 0), B_3(0, 0)\}$
PFR [12] (FFR-FI [8]) (Hard frequency reuse [33]) Figure 4	$\mathcal{B} = \{B_1, B_2, B_3, B_4\}$ $\mathcal{R} = \{r_1, r_2\}$ $\alpha = \{a_1, a_2\}$) $s_1 = \{B_1(a_1, r_1), B_2(0, 0), B_3(0, 0), B_4(a_2, r_2)\}$ $s_2 = \{B_1(a_1, r_1), B_2(0, 0), B_3(a_2, r_2), B_4(0, 0)\}$ $s_3 = \{B_1(a_1, r_1), B_2(a_2, r_2), B_3(0, 0), B_4(0, 0)\}$
PFR with only one interference in the worst case [22] Figure 5	$\mathcal{B} = \{B_1, B_2, B_3, B_4\}$ $\mathcal{R} = \{r_1, r_2\}$ $\alpha = \{a_1, a_2\}$ $s_1 = \{B_1(a_1, r_1), B_2(a_2, r_2), B_3(0, 0), B_4(a_2, r_2)\}$ $s_2 = \{B_1(a_1, r_1), B_2(0, 0), B_3(a_2, r_2), B_4(a_2, r_2)\}$ $s_3 = \{B_1(a_1, r_1), B_2(a_2, r_2), B_3(a_2, r_2), B_4(0, 0)\}$
SFR (Original) [9] Figure 6-(a)	$\mathcal{B} = \{B_1, B_2, B_3\}$ $\mathcal{R} = \{r_1, r_2\}$ $\alpha = \{a_1, a_2\}$ $s_1 = \{B_1(a_1, r_1), B_2(a_1, r_1), B_3(a_1, r_1) \cup B_3(a_2, r_2)\}$ $s_2 = \{B_1(a_1, r_1), B_2(a_1, r_1) \cup B_2(a_2, r_2), B_3(a_1, r_1)\}$ $s_3 = \{B_1(a_1, r_1) \cup B_1(a_2, r_2), B_2(a_1, r_1), B_3(a_1, r_1)\}$
SFR (2nd definition) [11] Figure 6-(b)	$\mathcal{B} = \{B_1, B_2, B_3\}$ $\mathcal{R} = \{r_1\}$ $\alpha = \{a_1, a_2\}$ $s_1 = \{B_1(a_1, r_1), B_2(a_1, r_1), B_3(a_2, r_2)\}$ $s_2 = \{B_1(a_1, r_1), B_2(a_2, r_2), B_3(a_1, r_1)\}$ $s_3 = \{B_1(a_2, r_2), B_2(a_1, r_1), B_3(a_1, r_1)\}$
SFFR [16] Figure 7	$\mathcal{B} = \{B_1, B_2, B_3, B_4\}$ $\mathcal{R} = \{r_1, r_2\}$ $\alpha = \{a_1, a_2\}$ $s_1 = \{B_1(a_1, r_1), B_2(a_1, r_1), B_3(a_1, r_1), B_4(a_2, r_2)\}$ $s_2 = \{B_1(a_1, r_1), B_2(a_1, r_1), B_3(a_2, r_2), B_4(a_1, r_1)\}$ $s_3 = \{B_1(a_1, r_1), B_2(a_2, r_2), B_3(a_1, r_1), B_4(a_1, r_1)\}$
FFR with multiple user class [21] Figure 8	$\mathcal{B} = \{B_1, B_2, B_3\}$ $\mathcal{R} = \{r_1, r_2, r_3\}$ $\alpha = \{a_1, a_2, a_3\}$ $s_1 = \{B_1(a_3, r_3), B_2(a_1, r_1), B_3(a_2, r_2)\}$ $s_2 = \{B_1(a_2, r_2), B_2(a_3, r_3), B_3(a_1, r_1)\}$ $s_3 = \{B_1(a_1, r_1), B_2(a_2, r_2), B_3(a_3, r_3)\}$
IFR [23] Figure 9	$\mathcal{B} = \{B_1, B_2, B_3\}$ $\mathcal{R} = \{r_1\}$ $\alpha = \{a_1\}$ $s_1 = \{B_1(a_1, r_1), B_2(a_1, r_1), B_3(a_1, r_1)\}$ $s_2 = \{B_2(a_1, r_1), B_3(a_1, r_1), B_1(a_1, r_1)\}$ $s_3 = \{B_3(a_1, r_1), B_1(a_1, r_1), B_2(a_1, r_1)\}$
EFFR [24] Figure 10	$\mathcal{B} = \{B_1, B_2, B_3, B_4, B_5, B_6\}$ $\mathcal{R} = \{r_1, r_2\}$ $\alpha = \{a_1, a_2\}$ $s_1 = \{B_1(a_2, r_2), B_2(a_1, r_1), B_3(0, 0),$ $B_4(a_1, r_1), B_5(0, 0), B_6(a_1, r_1)\}$ $s_2 = \{B_1(0, 0), B_2(a_1, r_1), B_3(a_2, r_2),$ $B_4(a_1, r_1), B_5(0, 0), B_6(a_1, r_1)\}$ $s_3 = \{B_1(0, 0), B_2(a_1, r_1), B_3(0, 0),$ $B_4(a_1, r_1), B_5(a_2, r_2), B_6(a_1, r_1)\}$

schemes [56]. The optimal settings for ICIC are particular to each geometry and the performance of static ICIC schemes is different from one network to another. Static ICIC schemes could penalize cells receiving more interference due to irregular cell layouts and thus irregular cell layouts require non-regular bandwidth allocations that suites each cell geometry. Furthermore, static schemes are unsuitable for Heterogeneous Networks (HetNets) with femto/pico/macro cells as these cells are placed at the end-user's locations in an ad-hoc manner making any prior frequency planning difficult [12]. These shortcomings of static schemes are addressed by dynamic ICIC schemes as they do not require prior frequency planning and operate based on dynamic interference information from surrounding transmitters.

III. DYNAMIC ICIC: CELL COORDINATION-BASED SCHEMES

The scale and complexity of modern mobile communication systems have motivated the exploration of *cell coordination-based schemes* as possible models for management and control

of such highly complex systems [61]. The complexity of these systems is due to several factors including the diversity of applications, volume of connections, geographic spread of users, localized ownership of the network, and "connectivity, anytime, anywhere" with an ever increasing demand for bandwidth.

The issue with the apriori frequency planning schemes discussed in Section II is that the inhomogeneous traffic load and varying user group distribution within each cell is ignored to simplify the cell-planning phase. This consequently leads to significant performance degradation in terms of cell and user throughput [33]. In realistic systems, the traffic load is unlikely to be spatially homogeneous and may exhibit significant variations over time. For example, one might see concentrations of users in different regions at different times of the day, e.g, train stations, shopping districts, and lunch time. As such, it is crucial that interference coordination schemes should be designed to adapt to different network interference conditions, user traffic load, and user distribution in order to maximize the total network throughput.

Cell coordination schemes have emerged as an efficient solution to cope with the continuous dynamic traffic load changes in cells. In cell coordination, interference reduction is realized by real time coordination using adaptive algorithms to efficiently manage the resource utilization among cells without apriori resource partitioning.

Although this solution presents a flexible framework as no apriori frequency planning is required, it may however require a signaling interface between different eNBs in order to achieve the required coordination which is considered as a serious complexity with respect to both overhead and delay. Various cell coordination-based schemes present tradeoffs between implementation complexity and the overhead of signaling. The problem of resource allocation with dynamic demand is known to be NP-hard [31]. Using an exact method is computationally inefficient as the problem involves extremely large search spaces with correspondingly large number of potential solutions. As the new resources configurations must be computed at run-time, computational efficiency is favored over model accuracy.

Due to the complexity of the dynamic ICIC problem, most of the performance evaluations are based on simulation models. A principal problem with simulation evaluations during comparing different schemes is the lack of common context, scenarios and evaluation metrics. Thus, unified realistic scenario data sets are needed that define common conditions such as cells layout, number of channels, propagation data and traffic intensity as well as a unified set of metrics to be used to evaluate various approaches. This issue has been open since 1996 [48]. There were some initiatives (e.g., [59]) to provide a realistic common context and scenario data sets. However, these initiatives did not receive enough attention in the community, and thus, the challenge of performance comparison for dynamic ICIC still remains. Accordingly, comparison of the permeance of various scheme may not be accurate enough, and hence, in this work, we base our review for the various schemes on the advantages and disadvantages, computational complexity, signalling overhead, and practicality of implementation.

In this section, a new classification model is presented and used to classify and explain various coordination-based schemes. The proposed classification model makes use of four dimensions, namely, the *optimization objective*, *power control technique*, *channel allocation recommender*, and the *fairness to UE*. In the following, the four classification dimensions are explained and then used to review key coordination-based schemes reported in the literature.

A. Classification of Coordination-based ICIC Schemes

Coordination-based schemes can be categorized, based on the level of coordination, into four main categorizes: *centralized*, *semi-distributed*, *coordinated-distributed*, and *autonomous-distributed*. In addition to these levels of coordination, various cell coordination-based schemes can be differentiated based on the following four dimensions:

1) *Optimization objective.* This dimension refers to the performance objective that needs to be optimized under a particular coordination-based scheme. Proposed schemes in the literature focus on one or more of the following optimization objectives:

- Maximize Throughput (T): The system attempts to dynamically find the assignment matrix (channels assigned to users) so that the total throughput is maximized.
- Minimize Interference (I): The system attempts to dynamically restrict interfering channels so that the number of interfering UE or the effect of interference is minimized.
- Minimize Power Usage (P): The system attempts to dynamically find the assignment matrix (channels assigned to users) so that the total power usage based on the interference levels reported by UEs for different channels is minimized.

2) *Power control technique.* Some studies in the literature (e.g., [29]- [31], [35], [40]) suggest that power control does not always yield significant performance gain in OFDM systems compared to the complexity it adds to the operations of the system. Accordingly, these studies adopt a simple *binary* power control (b, for short) model in which either a channel can be assigned to a UE with maximum power or not. However, some coordination based schemes (typically, autonomous distributed schemes) implement more sophisticated *full* power control (f) to allocate frequencies already in use by neighboring cells to UEs while minimizing the interference effect.

3) *Channel allocation recommender.* This dimension refers to the entity in the network that creates the *wish list* that contains possible channels to be assigned to each UE. Various schemes make use of different types of entities in order to create the wish list of each UE. Typically, a coordination-based scheme uses one of the following network element types to create the wish list for each UE:

- RNC: A central entity is responsible for finding a set of candidate channels for every UE such that there is no conflict between any combinations. Every eNB, on the other hand, assigns channels to UEs subjected to the constraints delivered by the central entity.

- eNB: Each eNB is responsible for creating a list of channels to use or restrict based on the information exchanged with the neighboring eNBs.
- UE: Each UE is responsible for collecting information of the various possible channels that can be used or need to be restricted. UE then sends this list to its eNB, which in return processes the lists obtained from all UEs and coordinate with the neighboring eNBs.

4) *Fairness to UE.* To ensure fairness among all UEs, some coordination-based schemes provide mechanisms to guarantee a minimum rate to all UEs. In such schemes, a minimum rate is placed for both interior and edge UE that the scheme guarantees. Some other coordination-based schemes adopt the best effort model, where the scheme attempts to achieve a required performance regardless whether the needs of all UEs are satisfied or not.

Based on the above dimensions, a given cell coordination-based scheme S can be defined using the notation $S = O_e^t(r_{min}^E, r_{min}^C, p)$, where:

- $O \in T, I, P$: an alphabetical set that describes the optimization problem the scheme is attempting to solve. Where T: Allocate channels to Maximize Throughput, I: Restrict channels to Minimize Interference, and P: Allocate channels to Minimize Power Usage.
- $e \in \{UE, eNB, RNC\}$ a subscript that represents the entity that creates the channels wish list.
- $t \in \{SD, CD, AD\}$ a superscript that represents the type of the scheme whether it is semi-distributed (SD), coordinated-distributed (CD), or autonomous-distributed (AD) scheme.
- r_{min}^E and r_{min}^C indicate the scheme fairness by representing the minimum guaranteed rate for edge and central UE respectively, where: $r_{min}^E, r_{min}^C \in \{0, r_1, r_2\}$.
- p indicates the possible fractions of the maximum power that can be allocated to a channel. $p \in \{\{b : b \in \{0, 1\}\}, \{f : f \in \mathbb{Q}, 0 \le f \le 1\}\}$

Table II summarizes the various cell coordination-based schemes using the above classification parameters.

In the following subsections, different coordination-based schemes are grouped based on the coordination category (*centralized*, *semi-distributed*, *coordinated-distributed*, or *autonomous-distributed*). Under each category, various schemes are explained based on the nations described above.

B. Centralized Schemes

In *centralized schemes*, a central control unit collects all the channel state information (CSI) of every UE in the system and allocates available RBs to each eNB trying to maximize the capacity according to fairness and power constraints. Therefore, each eNB has to forward the received CSIs of each UE to the centralized controller and receive back the allocation information before transmitting, resulting in a high backhaul signaling.

However, without an efficient and fast infrastructure, centralized scheduling is a hard task due to the stringent time required to exchange the inter-cell scheduling information and the large feedback required by the UEs to send all the CSIs [52]. For these reasons emerging cellular networks such as the LTE-Advanced systems have eliminated the central

TABLE II
SUMMARY OF CELL COORDINATION-BASED SCHEMES.

Representation	Example Scheme(s)
$S = I_{eNB}^{SD}(r_1, r_2, b)$	[12] [29]
$S = T_{RNC}^{SD}(0, 0, b)$	[30] [31]
$S = I_{UE}^{CD}(r_1, r_2, f)$	[32] [38]
$S = P_{eNB}^{CD}(r_1, 0, f) + T_{eNB}^{CD}(0, r_2, f)$	[33]
$S = T_{eNB}^{CD}(r_1, r_2, f)$	[34] [36] [37]
$S = T_{eNB}^{CD}(r_1, r_2, b)$	[35]
$S = T_{eNB}^{CD}(0, 0, f)$	[39]
$S = T_{eNB}^{CD}(0, 0, b)$	[40]
$S = P_{eNB}^{AD}(0, 0, f)$	[43] [44]
$S = P_{UE}^{AD}(r_1, r_2, f)$	[45]
$S = P_{eNB}^{AD}(r_1, r_2, f)$	[46]
$S = I_{UE}^{AD}(0, 0, f)$	[47]

control unit and relied on inter-eNB coordination over the X2 interface with no central coordinator in a flat architecture [50]. Examples of centralized scheme can be found in [26]-[28]. It is worth noting that, most of the centralized schemes were designed for the Time Division Multiple Access (TDMA) and Code Division Multiple Access (CDMA) systems.

C. Semi-Distributed Schemes

As the name suggests, semi-distributed schemes (e.g., [12] [29]- [31]) are neither fully centralized nor fully distributed. Coordination in these schemes is typically performed at two levels: the central entity level and the eNBs level. Similar to centralized schemes, semi-distributed schemes implement a central controlling entity that controls a number of eNBs. However, semi-distributed schemes make use of the central entity to allocate in each super-frame a bulk of resources to each eNB instead of allocating the channels directly to each UE on a frame bases as in the centralized schemes. Accordingly, in semi-distributed schemes, each eNB is responsible for allocating channels on the frame level to the UEs that it serves.

Similar to the case of centralized schemes, efficient and fast infrastructure is still required, but a bit less complex, in order to exchange the inter-cell scheduling information and feedback on the backhaul link. However, the resource allocation problem is distributed between the central entity and eNBs, which in turn can reduce the computational complexity of the overall scheme.

It is worth noting that, in homogeneous networks with no central controllers, semi-distributed schemes, similar to the centralized schemes, are not practical for implementation [12]. However, the semi-distributed approach can still be
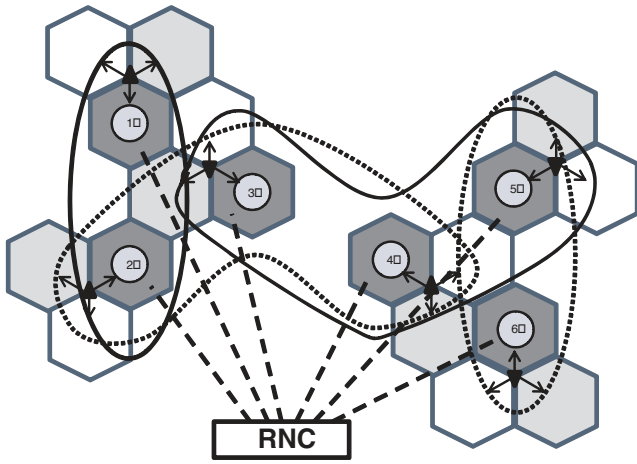
Fig. 12. Concept of Interferer groups used in schemes of Ref. [12] and Ref. [29]. Shown interferer groups are: $\{1, 2\}, \{2, 3, 4\}, \{3, 4, 5\}$, and $\{5, 6\}$.

used for enhanced ICIC (eICIC) in heterogeneous networks (HetNets) with femto/pico/macro cell layouts [60]. The full study of HetNets and eICIC schemes is out of scope of this paper.

In the following, we summarize some of the semi-distributed schemes reported in the literature. The schemes are grouped based on the classification dimensions explained in Section III-A.

1) $S = I_{eNB}^{SD}(r_1, r_2, b)$ Schemes: In [12] and [29] Rahman *et al.* followed a heuristic approach to solve the resource allocation problem in a multi-cell system. In their approach, each eNB detects if a downlink transmission from neighboring eNBs would be a potential cause of dominant interference to its UEs. To do so, each UE sends the channel state information (CSI), including information on the two most dominant interference received from neighboring eNBs, to the serving eNB. Based on the dominant interference information collected, each eNB forms an interference group (Figure 12) with those reported eNBs that cause dominant interference. Each eNB then creates a *wish list* of RBs to be restricted in its neighboring cells, and sends this list to the central entity along with the utility measure of the RBs in the requesting cell via the X2 interface. This central entity can be a RNC, or a mobility management entity (MME).

The central controller gathers all such requests, processes them to prepare a refined list of RBs restrictions to be applied in all involved eNBs, and sends the restriction decision to all UEs. This restriction process is refreshed from time to time within an interval that is shorter than the channel coherence time, which depends on the mobility of UEs. The decision of the central entity on a RB under consideration reported by a given eNB could be one of the following: (1) other eNBs should mask the dominant interference frequencies, (2) other eNBs should keep using the dominant interference frequencies but the reporting eNB should mask them, or (3) all eNBs should use the dominant interference frequencies concurrently as the mutual interference permits to achieve better aggregate spectral efficiency.

In [29], Rahman *et al.* used equal power allocation to all RBs and modeled the ICIC problem using binary integer linear programming (ILP) that can be entirely solved at the central

entity. Since the complexity of the ILP increases exponentially with the number of variables and constrains, the problem is solved by iteratively solving a set of sub-problems of allocating subsets of unassigned channels to subsets of rate unsatisfied UEs. The overall computational complexity of this approach is $O(min(K, J) \times N_s^4 \times M_s)$, where K and J are the number of UEs and RBs subsets, respectively, and M_s and N_s are the number of UEs and RBs in each subset, respectively.

On the other hand, in [12], Rahman *et al.* divided the computations to be shared between the central entity and the eNBs. The algorithm at the eNB uses the iterative Hungarian algorithm in order to process the RB restriction requests of the UE, and generate the wish list of RB restrictions. The wish list is then forwarded to the central entity that solves the restriction requests in an optimal manner and returns the decision to the involved eNBs to apply it locally. Two RB restriction approaches were proposed in this scheme: (1) dropping the RB, or (2) using the RB but with 10 dB lower power. The overall complexity of the cell-level algorithm is $O(M \times N + (min(M, N))^2 \times max(M, N))$, which is dominated by the complexity of the Hungarian algorithm, where M and N are the number of UEs and RBs in a cell, respectively. Whereas the complexity of the algorithm at the central controller is $O(L \times N_r)$, where L is the number of cells that have conflicting restriction requests and N_r is the number of RBs under consideration for restriction. The results of the performance evaluation of this scheme can be summarized as follows:

- Higher throughput can be achieved by restricting more and more RBs which implies more penalties to neighboring cells. However, substantial gain can be achieved by restricting only the most dominant interferer, and thus, only this interferer should be restricted. Only in the case of severely rate deprived UEs, the two most dominant interferers can be restricted.

- An average of 12.5% to 20% of RBs need to be restricted in each cell in order to obtain high throughput gains. However, this large number of RB restrictions in each cell causes a considerable loss in resource, and thus, only justifiable restrictions should be made. To that end, the following two restriction policies were evaluated in [12]: (1) RB restrictions are made only in favor of UEs that have received less than the average service in the cell, and (2) RB restrictions are also made for the UEs with good service status only when a considerable gain can be achieved. The two policies represent a trade-off between performance and signaling overhead complexity as the latter provides better performance at the cost of an increased number of restricted RBs and higher signaling overhead. This is due to the fact that in the second policy all UEs need to forward information of the two most dominant interferers to the serving eNB and not only the rate deprived UEs as in the first policy.

- Restricting RBs by totally dropping the RBs achieves performance for edge-UE that is comparable to the reuse-3 scheme, while at the same time maintains a lower cell throughput and delay performance as compared to the reuse 1 scheme. On the other hand, restricting RBs by only lowering the used power by 10 dB maintains the cell

throughput and delay performance at levels similar to that of the reuse 1 scheme, but at the cost of a reduced benefit to edge-UEs.

Although, a central entity is required for the resolution of the conflicting requests, the algorithm can be applied to homogenous networks without a central controller (eg., 3GPP LTE and LTE-Advanced networks) with expected degradation in performance as conflict resolution is expected to be sub-optimal. In this case, resolutions can be performed through negotiations among neighboring eNBs using X2 interface that inter-connects eNBs. If two eNBs wish to restrict a RB to each other, the decision should result in favor of the eNB that foresees higher utilization on that RB. Accordingly, the proposed scheme can be classified as a coordinated-distributed scheme.

2) $S = T_{RNC}^{SD}(0, 0, b)$ *Schemes:* In [30], Li *et al.* presented a semi-distributed scheme that reduces the overhead and computational load required to exchange all channel state information (CSI) of UEs and traffic status information between the eNBs and the central entity at the frame level. The scheme proposes splitting the resource allocation decision between the central entity and eNBs. At super-frame level, the central entity makes decision on which RB is used by which eNB and recommends a UE to be assigned to the RB to maximize the system throughput and minimize ICI. Then, at the frame-level, eNBs make the actual pairing between the RB and the UE. For a given particular frame, if the UE recommended by the central entity has traffic to send, the eNB will agree with the recommendation of the central entity; Otherwise, if this UE has no traffic to send in any of the frames of the super-frame, the eNB will make its own decision based on the traffic (buffer occupancies) of the UE and fading channel conditions. As a result, the RB is always assigned to the UE with the highest utility value, which is a function of both channel and traffic conditions. Given that at each super-frame, the interference from neighboring eNBs to UEs is pre-determined by the *RBs to eNBs allocation* made by the central entity, thus, re-allocating a RB to a UE with traffic to send instead of the recommended UE with no traffic, will always result in an improved throughput for the eNB.

According to the above discussion, it appears that the decision algorithm of the central entity performs interference avoidance, whereas the decision algorithm of the eNB performs channel/traffic adaptation. Specifically, the central entity will be dedicated to coordinate the mutual interference between cells, which will reduce the information update rate between the central entity and the eNBs to a super-frame level. The eNBs, on the other hand, will make real time decisions on RB assignment at the user packet level (frame level). As a result, both the mutual interference diversity and the fading channel/bursty traffic diversity are efficiently exploited. However, one of the main disadvantages of this scheme is the allocation of an equal power to all RBs to simplify the central entity algorithm. Manipulating the RB allocated power can achieve higher spectral efficiency by allocating the same RB in different cells at different power levels, and lower ICI by minimizing the power levels of the dominating interferers.

The central entity algorithm uses a greedy approach to assign a RB to the UE that has the highest system throughput
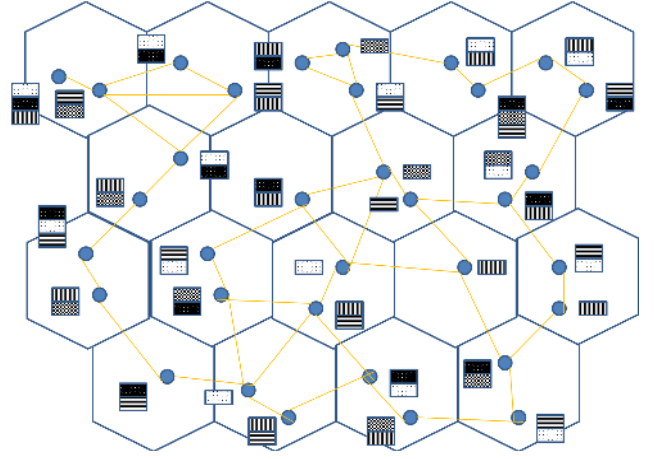

Fig. 13. Graph coloring of the outer optimization problem used in scheme of Ref. [31].

improvement value within a certain eNB. The *RBs to eNBs assignment* is progressively performed to provide the most improvement to the system throughput. If none of the UEs in an eNB has a positive system throughput improvement value, then the RB is not assigned to this particular eNB. Throughput ratios (proportional to the number of active users) are set among the cells to prevent edge-UE starvation. Also, the RB assignment always starts with the eNB that is most under-assigned and ends with the eNB that is most over-assigned as the order with which eNBs are evaluated has a considerable impact on the overall final throughput. The first eNB will enjoy the use of a "clear" channel, whereas, the rest of eNBs will suffer from the interference resulted from previous assignments. When no more eNBs can increase the overall throughput of the network by using a particular RB, the RB is not assigned to any additional eNBs. Finally, the eNB algorithm attempts to find the UE that has the highest throughput for each RB.

The algorithms discussed above perform the needed operations for both interference avoidance and traffic adaptation with a linear-complexity with respect to the number of users and channels. The computational complexities of the central entity and the eNB algorithm are, respectively, $O(N \times L \times M_{eNB})$ and $O(N_{eNB} \times M_{eNB})$. Where N is the number of RBs, L is the number of eNBs, M_{eNB} is the number of users in an eNB, and N_{eNB} is the number of RBs assigned by the RNC to an eNB. It is worth pointing that, since the computational complexity of the central entity algorithm is a function of the number of the eNBs, this algorithm may not operate efficiently for systems that has small cells with a large number of interfering eNBs.

In [31], Neckera *et al.* split the RB allocation problem into an *outer* and *inner* optimization problems as follows.

- *The outer optimization problem* is solved at the central entity and is modeled as a global interference graph based on the global UEs information collected from all eNBs. In this graph, the vertices represent the UEs, and the edges represent critical interference relations between them. The goal is to find a set of colors (set of RBs) for every UE

such that there is no conflict between any combinations of colors in the sets. This problem is known as fractional graph coloring and can be briefly explained as follows. First, the graph is constructed by evaluating the interference that a transmission to one UE causes to any other UE. For each UE, the total interference that the UE receives from other eNBs is calculated. The largest interferers are then blocked from using the same set of RBs by establishing a relation in the interference graph. Resources need to be assigned to UEs such that no two UEs are assigned the same resources if they are connected in the graph.

To solve the outer optimization problem, first the graph is colored by means of the sub-optimal heuristic Dsatur (sequential graph coloring algorithm heuristic). Next, all UE are traversed in a random order and assigned a second color where possible. The process is then repeated until no more extra colors can be assigned to any UE as illustrated in Figure 13. Color assignment is based on that the largest interferers are blocked from using the same RBs. RBs need to be assigned to the UEs such that no two UE are assigned the same RB if they are connected in the graph. The computational complexity of the outer optimization algorithm is $O(M^2 + M \times N)$, where M is the number of UEs in all cells and N is the number of RBs.

- *The inner optimization problem* is solved locally at each eNB and is modeled as a local interference graph based on the local state information and subject to the allocation constraints delivered by the outer optimization problem, which is solved at the central coordinator and results are communicated to the eNB at regular time intervals (in the order of seconds). The goal of the inner optimization problem is to coordinate the transmissions in a cell by assigning every UE to one or more RB. However, those RBs must be chosen from the set of RBs assigned to this UE by the central entity.

The inner optimization problem is then solved by means of genetic algorithms. Initially, all UE are listed along with their RBs that were assigned during the outer optimization. A placement algorithm traverses the list and assigns the first possible and free RB to the UEs. The RBs must not yet be occupied, and the assignment must not be in conflict with the inner interference graph. During the subsequent iterations, possible solutions (assignment matrices) are evaluated with the objective of maximizing the number of scheduled UEs.

The effect of the number of possible solutions in the pool (P) and the number of iterations/generations (G) on the edge-UE performance was studied in [31]. It was found that at small G, edge-UEs will most likely be not served, thus decreasing the edge-UEs performance. The aggregate throughput however, depends much less on the G than the edge-UE throughput. The same also holds for the number of possible solutions in the pool P. It was observed that as few as $G= 10$ generations bring the number of unserved UEs below one. For $G= 100$, the algorithm comes close to its optimum performance. The best performance for a small computational effort can be achieved for $P \approx 2 \times G$. The inner optimization algorithm (based on genetic algorithms) computational complexity is $O(P \times G)$.

From the above, it appears that the complexity of the *inner optimization algorithm* is low as the genetic algorithm converges after few number of generations. This allows for efficient hardware-based real-time implementations of the algorithm. However, the computational complexity of the *outer optimization algorithm* is function of the square of the number of all UEs in the system, which makes it difficult for this algorithm to converge in timely manner for networks with large number of UEs.

Similar to work reported in [30], Neckera *et al.* in [31] allocate equal power to all RBs to simplify the central entity algorithm which minimizes the resources utilization. However, manipulating the RB allocated power can achieve higher spectral efficiency by allocating the same RB in different cells at different power levels, and lower the ICI by minimizing the power levels of the dominating interferers.

D. Coordinated-Distributed Schemes

Frequent RB allocation updates may lead to a higher throughput as they allow for capturing more accurate interference avoidance gain and partial fading diversity gain information [30]. As a result, a better decision-making process can be performed especially in the emerging systems that support high user mobility. *Centralized and Semi-distributed schemes* are often too heavy to be implemented as all the interference information on all RBs have to be gathered at a central controller. In reality, the amount of information needed from eNBs to the central entity can be prohibitively large [30]. Consequently, the frequency of exchanging information with the central entity must be reduced, which in turn may degrade the overall performance of the system. A scheme needs to make a trade-off between increasing coordination between eNBs, to increase the overall system throughput; and reducing the cost of backhaul communication and intra-node processing [49].

In *coordinated distributed schemes*, resource allocation is performed only at the eNB level, without the need of a central entity to perform the coordination. However, coordination is still needed between eNBs to exchange CSI reports in order to perform global ICIC. Such coordination must be taken into consideration when designing *coordinated distributed schemes*, especially in very fast fading environments due to the fact that the LTE X2 interface used for coordination has a non-negligible latency [45].

Coordinated distributed schemes have a clear advantage over the *semi-distributed schemes* presented above. Such advantages include reducing time and signaling overhead resulted from regular communication between eNBs and the central entity, and reducing the network infrastructure complexity by eliminating the central entity. Thus, *coordinated distributed schemes* are more attractive for practical implementation. However, the realization of these schemes has remained limited largely due to constraints on inter-eNB communication and the latencies involved in information exchange for distributed eNBs [53].

Several Coordinated distributed schemes have been proposed in the literature (e.g., [32]- [40]). Schemes under this category typically partition the complex multi-cell optimiza-

tion problem into distributed single-cell optimization problems. These single-cell optimization problems can be solved by each individual eNB using local information from the served UEs while exchanging minimal amount of information required for global ICIC with neighboring eNBs. It worth noting that, some of the schemes under this category make an *implicit* use of some of the concepts in the frequency reuse schemes discussed in Section II. For example, the schemes proposed in [32]- [35] adopt the power constraints of the soft frequency reuse (SFR) scheme.

In the following, we summarize some of the coordinated-distributed schemes reported in the literature. The schemes are grouped based on the classification dimensions introduced in Section III-A.

1) $S = I_{UE}^{CD}(r_1, r_2, f)$ Schemes: In [32], Triki *et al.* proposed the Exchange-based Interference Coordination (EIC) algorithm based on resource exchange and borrowing between eNBs. The proposed algorithm runs on the eNBs with no need for a central coordinator. It depends entirely on the local information from the served UEs and the information exchanged between neighboring eNBs. The EIC algorithm is performed in the following three phases:

- *Initial phase:* in this phase the available bandwidth is divided into three non-overlapping sub-bands and each sub-band is assigned to an eNB making a cluster of three eNBs. In other words, the scheme starts with a reuse-3 configuration. The authors define the *owner eNB of a RB* as the eNB that has this RB in the sub-band that was initially assigned to this eNB.

- *Proactive phase:* in this phase, eNBs starts borrowing RBs from their neighboring eNBs if the owned RBs cannot satisfy the minimum rates of their UEs. The authors define the *borrowing eNB of a RB* as the eNB that borrows this RB from a neighboring eNB (i.e, the requested RB is not in the sub-band initially assigned to the borrowing eNB). The following activities occur during the borrowing process:

  i Each eNB evaluates the quality of its RBs and sends a Load Information Message (LIM) to its neighbors. The LIM contains the status of each and every RB of the eNB, including whether the RB is allocated or not and the level of interference in this RB (high, medium, low, or clear). The level of interference value is determined using a set of predefined thresholds. One of the weak points of this algorithm is that no method or study is provided to optimize the values of these thresholds and understand their effect on the overall performance.

  ii A borrowing eNB, which desires to borrow a RB outside its sub-band, chooses the best RB among those available in the LIM. The utilization of this RB by the *borrowing eNB* must not cause harmful interference to the *owner eNB* of this RB. This check can be performed since each eNB knows the assigned sub-band for every neighboring eNB. To prevent harmful interference to the *owner eNB* due to RB borrowing, the *borrowing eNB* compares the interference level of all "borrower-able" RBs on the *owner eNB* to a pre-defined borrowing threshold. If RB interference level is inferior to the borrowing threshold, the eNB can borrow the RB without further

information of the RB *owner eNB*. While taking the borrowing decision based on a static threshold is simple, the authors state that the choice of other borrowing decision mechanisms may improve the performance of the proposed algorithm, without necessary changing the operational concept of the algorithm.

- *Reactive phase:* in this phase the *owner eNBs* exchange the interference level of the borrowed RBs belonging originally to their sub-band. According to interference level of the RB, *borrowing eNBs* adapt their transmission power used for the borrowed RB. The *borrowing eNB* reduces the transmission power of a borrowed RB by 25% if it receives high interference indicator for this RB from the *owner eNB* for the first time. If the *borrowing eNB* receives a second high interference indicator, it reduces again the transmission power by another 25%. If the level of interference remains high and the *borrowing eNB* receives a third high interference indicator, the *borrowing eNB* releases immediately the RB to avoid causing higher interference to the *owner eNB*.

The proposed *EIC* algorithm runs periodically, where each period is triggered by a *proactive phase* (RBs borrowing) followed by three successive *reactive phases* (borrowed RBs power adjustments). At the beginning of a *proactive phase*, the eNB releases the borrowed RBs that suffer from a high interference and keeps the other borrowed RBs and their transmission power if they are still needed.

The performance evaluation of the EIC algorithm showed that it outperforms static SFR in terms of throughput and number of served users, while utilization a lower number of resources. This high spectral efficiency ($\frac{bits}{RB}$) of the EIC algorithm is attributed to the efficient selection of RBs during the *proactive phase* and the efficient power management of RBs at the *reactive phase*.

The computational complexity of the *proactive phase* (RBs borrowing) is $O(M \times L \times N)$, where M is the number of unsatisfied UE in a cell, L is the number of neighboring eNBs, and N is the number of RBs available for borrowing from a neighboring eNB. The computational complexity of the *reactive phase* (borrowed RBs power adjustments) is $O(N)$. As the available bandwidth is divided into three non-overlapping sub-bands and each sub-band is assigned to an eNB, the number of RBs available for borrow from a neighboring eNB N represents the whole neighboring sub-band, which is third of the available bandwidth.

Zheng *et al.* in [38] proposed a novel distributed QoS-guaranteed ICIC scheme. This scheme is based on allowing edge UEs in an over-loaded cell to send to the neighboring eNBs a request of restraining them from using the same RBs. Thus, under-loaded cells make concessions to the requesters and shoulder part of the heavy burden of the over-loaded cell in a new form of load balancing. The proposed scheme contains the following two procedures, both of which are performed by the eNB on two different time levels:

- The *intra-cell fast scheduling algorithm based on the Hungarian algorithm* is performed on the frame level to assign RBs to UEs to meet their QoS targets. The algorithm decisions are based on the instantaneous (frame level CQI reports) SINR of UEs considering the instantaneous channel

gain which is related to flat fading and fast fading. During this process, certain under-served UEs have the opportunity to prepare a request of limiting the use of some of their currently scheduled RBs in their dominant interfering cells. The authors define an under-served UE as a UE with the ratio of its target data rate to its average data rate greater than unity. The complexity of this *intra-cell fast scheduling* algorithm is $O((M \times N)^2)$, where M and N are the numbers of the UEs and RBs in the cell, respectively. Although the intra-cell scheduling takes into account the UEs' fairness, it cannot resolve the problem of severe throughput degradation of edge UEs when the cell is over- loaded as this problem requires actions form other eNBs. Thus, this *inter-cell interference coordination algorithm* is used to achieve global ICIC.

- The *inter-cell interference coordination (ICIC) algorithm based on X2 interface* is performed on the super-frame level to require other eNBs to reduce the transmitting power on certain dominant interference RBs, or to request not to use them at all based on the UE's restriction lists. The coordination among eNBs needs backhaul communication and signaling overhead. However, only medium-to-long term interactions between eNBs through the X2 interface are required. The eNBs employ the ICIC algorithm on a super-frame level to limit the usage of some RBs in order to minimize interference for over-loaded cells. The algorithm decisions are based on the long term (super-frame level CQI reports) SINR considering long term channel gain, which is related to large scale path loss and shadow fading. This algorithm consists of the following three operations:

  i *Over-loaded eNBs check eligibility of the restriction request*: Restricting RBs causes degradation in the spectral efficiency, hence the authors proposed the *Integral based Eligible UE Selection (I-EUS)* algorithm to choose which UE's restriction requests to be sent to the dominant interferers. The I-EUS algorithm considers a the restriction request if an UE eligible if the averaged ratio over the super-frame of the UE's target data rate to its average data rate exceeds a pre-defined threshold. The method of choosing the value for this threshold and the effect of this value on the algorithm performance were not discussed in [38]. A key characteristic of the I-EUS algorithm is that it takes into account the depth of hunger of under-served UEs in terms of QoS (under-served in a severe degree for a short period of time) as well as in the time domain (under-served in a lesser degree but for a long period of time). Requests under both cases will be accepted as the decision is based on the averaged ratio over the super-frame. Furthermore, from the perspective of load balancing, the issuing of eligible requests can be regarded as an indicator that the cell is over-loaded. The complexity of the *I-EUS* algorithm is $O(L \times M)$, where L is the number of the eNBs, and M is the number of the UEs in the cell.

  ii *Over-loaded eNBs prepare and send eligible restriction requests to the dominant interferers*: For eligible UEs, the eNB begins sending restriction requests to the dominant interferers starting by the worst RB to the less worst

RB, until the UE's target data rate is met. According to the eligible UE's channel condition on each of those RBs, the UE can either require the dominant interferer to reduce the transmitting power on the RB or force it not to use the RB. To make the decision, the authors proposed to compare the achievable rate of the UE, in case the dominant interferer reduced the transmitting power on the RB, to a pre-defined threshold. If the achievable rate of the UE was larger than the threshold, then only power reduction request is sent; Otherwise, a restriction request is sent. However, the selection of the value of the threshold and its impact on the performance of the algorithm were not studied in [38]. The complexity of this step is $O(L \times M_e \times N_u)$, where L is the number of the eNBs, M_e is the number of the eligible UEs in the cell, and N_u is the number of RBs scheduled for the UE.

  iii *Under-loaded eNBs resolve the received requests*: Since each over-loaded eNB compiles the list of restriction request independently, problem remains at the under-loaded eNBs (eNBs that receives restriction requests) to make a decision on which RBs to restrict, which to reduce their power, and which to leave as is. For instance, an under-loaded eNB might receive, for the same RB, two restriction requests from two of its neighbors and three power reduction requests from three other neighbors. For each RB, the proposed ICIC scheme groups eNBs requests of similar restriction type (e.g., total restriction or power reduction) and calculates the expected rate gains if this restriction type is applied on the RB. Accordingly, the restriction type that maximizes the rate gain is selected. The computational complexity of this process is $O(L \times N)$, where L is the number of eNBs, and N is the number of RBs.

Once the ICIC algorithm determines the permitted RBs for each cell, certain RBs may be unavailable to UEs, or certain RBs become particularly good to select. Then, the UE level resource allocation is performed by the fast scheduling algorithm in the next super-frame.

To reduce the computational and signaling complexity of the scheme above, authors suggested to limit the coordination to be within 7 cells only, making the algorithm suboptimal, yet practical. Even with this limitation, the computational complexity of this scheme is still relatively high compared to other coordinated-distributed schemes. This scheme follows the same strategy of restricting RBs presented in [12], but without the use of a central entity, which makes the scheme more practical for homogenous networks, and reduces the overall infrastructure complexity, and hence cost. It is worth pointing that this scheme, unlike the scheme proposed in [12], exploits full power control, which leads to a higher spectral efficiency.

2) $S = P_{eNB}^{CD}(r_1, 0, f) + T_{eNB}^{CD}(0, r_2, f)$ *Schemes:* In [33], Quek *et al.* combined the conventional SFR scheme [9] [Figure 6-(a)] with an adaptive frequency and power resource allocation approach. In this scheme, the SFR scheme is first used to assign channels to eNBs, and each eNB then uses heuristics and suboptimal algorithms to allocate power and

RBs to active UEs to guarantee a minimum achievable rate to all edge UEs. Unlike the previous schemes that divide the problem based on the timing level (super-frame and frame), this scheme divides the power and RB allocation problem to two sub-problems for edge and center UEs, respectively. Both problems are solved on the frame level using the instantaneous received SINR. A geometry-based approach was adopted for user group partitioning, where the edge and center UEs are differentiated based on their distances from the serving eNB. The following is a brief description of the two algorithms used for edge and center UEs.

- *Sum Power Minimization* algorithm for edge UEs: The first sub-problem is to allocate RBs to all edge UE to achieve a pre-defined minimum rate. Edge UEs have low SINR due to presence of ICI and significant path-loss. Thus, the most efficient way to improve their rate is to allocate more power to increase the SINR instead of increasing the bandwidth. This is the main idea of the *Sum Power Minimization Algorithm for edge UEs*. After the fixed sub-channel allocation using SFR, the algorithm starts with a single RB as the minimum number of RBs to be allocated to an edge UE and assumes fixed and equal power on all sub-channels. The algorithm iteratively solves the sum power minimization problem to allocate power to the different sub-channels allocated to edge UEs without exceeding the total transmission power of the eNB and subjected to a minimum rate constraint where each edge UE must achieve a pre-defined minimum rate. The algorithm ends by defining the set of edge sub-channels. The rest of the sub-channels are available for center UEs. Each eNB then shares the local sub-channel and power allocation with other eNBs. eNBs use the information exchanged between them to adapt the SFR planning to the traffic conditions. The computational complexity of this algorithm is $O(M_e \times N_e^4)$, where M_e is the number of edge UE, and N_e is the number of edge sub-channels.

- *Weighted Sum Rate Maximization* algorithm for center UE: The second sub-problem is to use the remaining sub-channels to maximize the throughput of the cell-center UEs. Center UEs generally have high SINR since they are closer to the serving eNB and faraway from the interfering eNBs. Thus, the most effective way to improve the rate of these UEs is to allocate more bandwidth instead of power. The proposed algorithm implements this idea by first uniformly allocating the residual transmit power over the remaining sub-channels, and then allocating the remaining sub-channels to the center UEs with the aim of maximizing the sum rate. The computational complexity of this algorithm is $O((M_c \times N_c)^2)$, where M_c is the number of center UEs, and N_c is the number of center sub-channels.

While edge UEs suffers the most, and thus, they were the main (and in most cases, the *only*) focus of most dynamic schemes, maximizing the rate of center UEs leads to a higher system utilization, and thus, it is important to optimize the center UEs allocations. Accordingly, unlike most dynamic schemes, this scheme considers both center and edge UEs while planning power and RB allocations. In addition, the scheme attempts to guarantee fairness to the edge UEs by
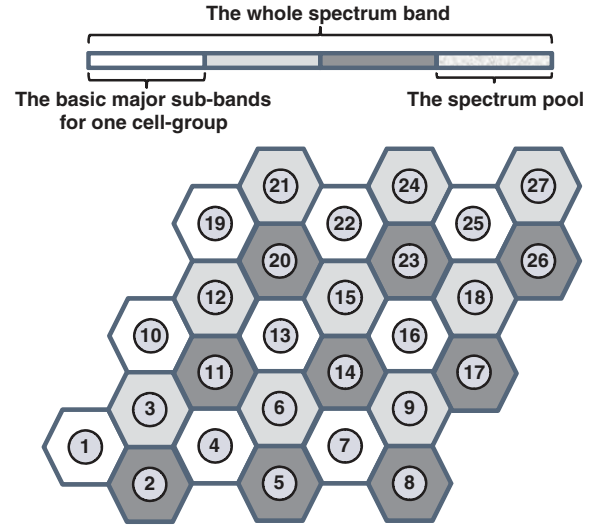

Fig. 14. Spectrum partitioning and spectrum pool used in the scheme proposed in Ref. [34].

ensuring that a minimum target rate is *pre-defined* and *fixed* for all edge UEs. However, it should be noted that in this scheme, no minimum rate is set for center UEs, and hence, their sub-channel allocation is based on best effort, which may adversely impact the fairness among center UEs.

A disadvantage of this scheme is the use of the static SFR scheme to initially divide the sub-channels among edge and center channels as this may limit the dynamic nature of the proposed scheme. Another disadvantage is the relatively high computational complexity and amount of information, which include all edge RB and power allocation information, that need to be exchanged between eNBs.

3) $S = T_{eNB}^{CD}(r_1, r_2, f)$ Schemes: Kun *et al.* proposed in [34], a distributed semi-static heuristic ICIC scheme based on the resource allocation rules of SFR in which the spectrum is divided into four partitions as illustrated in Figure 14. Each three neighboring cells are grouped to form a cluster, and each cell is assigned one partition to be used only by it as its major sub-band. The fourth partition acts as a spectrum pool and is divided into three smaller partitions. Each partition is assigned a priority to one of the cells in the cluster, but it can be used by any of the three cells in the cluster based on the loads within each cell. Each eNB calculates the average feasible data rate of the center UE, the current amount of the edge UE load, and the requested number of major sub-bands. Each eNB then reports its self evaluation result to its neighbor cells in the cluster, and accordingly, receives their evaluation information. Based on the list of evaluation results, each eNB decides on amount of the spectrum resources to utilize from the shared pool, if any. An eNB can use the portion of the spectrum pool in which it has the highest priority, and in addition, it can decide to use portions of which it has a low priority. This, however, can be done only if the cell has high traffic loads and the neighboring cell gives up its right in using this partition. The main disadvantage of this scheme is that eNBs can't utilize the full spectrum as for each cell there is a dedicated sub-band. In addition, the scheme utilizes the spectrum pool
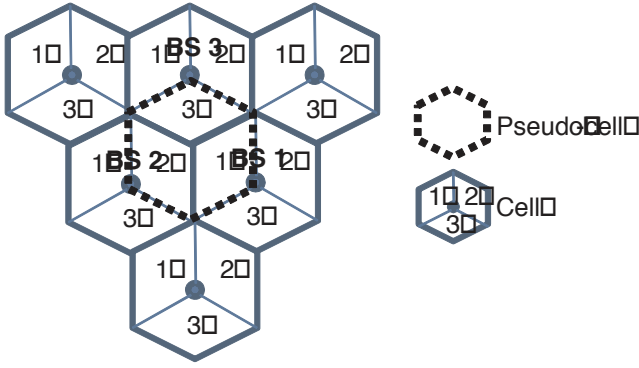
Fig. 15. Structure of the pseudo-cell concept used in the scheme proposed in [35].

at the *partition* granularity and not at the *channel*, which may result in inefficient utilization of the spectrum as the cell utilizing the partition may not require all the channels within that partition. The computational complexity of this scheme to allocate partitions from the spectrum pool to cells is $O(M_e + P \times L)$, where M_e is the number of edge UE, P is the number of partitions in the spectrum pool, and L is the number of neighboring cells in which message exchange occurs. It should be noted that the Proportional fairness (PF) algorithm is used in this scheme in order to schedule edge and center UEs to the channels available to the cell, which adds more computational complexity to the overall scheme.

Fraimis *et al.* presented in [36] and [37] an algorithm based on exchanging channel occupancy information (used channels) between eNBs to allow each eNB to individually allocate channels to its UEs such that the best value of channel gain is obtained. The objective of this scheme is to maximize the throughput of the cell-edge users in a decentralized manner with only minimal coordination between the neighboring eNBs. The scheme works as follows. The algorithm first attempts to find, using the exchanged information, a channel that is not used by an UE in any of the adjacent cells to allocate to edge UEs. The eNB tries to find the least used channels by the neighbors and assign them to the edge UEs. The process continues until all edge UEs are assigned channels or there are no more available channels. Next, the scheme starts allocating equal power to all edge UEs. One of the limitations of this scheme is that it assumes that an UE can be allocated a single channel only. This is assumption is not practical as it limits the achievable rate of an UE to the maximum achievable rate of best channel it can get. Another limitation in this scheme is the equal power allocation to all edge-users. This is not practical as edge users may be allocated channels that are used by neighboring cells, while others may be allocated interference-free channels. Accordingly, the former users may require higher power than the latter. The computation complexity of this scheme is $O(M_E \times N)$, where M_E is the number of edge UEs, and N is the number of channels.

4) $S = T_{eNB}^{CD}(r_1, r_2, b)$ *Schemes:* In [35], Kwon *et al.* proposed a scheme that attempts to minimize the maximum value of the QoS violation ratios in each sector. QoS violation ratio is defined as the ratio of the number of users whose rate requirements are not satisfied to the total number of users.

Assuming that the SINR value is identical for all the sub-carriers belonging to one sub-carrier group (SG) or channel, users feed back the channel condition on SG level to decrease the amount of feedback information. The system attempts to compute the value of the frequency reuse factor (FRF) of SGs dynamically and distribute the sub-carriers with high FRFs differentially among the neighboring sectors according to their load conditions. In this scheme, resource allocation is managed independently at each "pseudo-cell" composed of the major-interfering sectors belonging to the neighboring cells (See Figure 15) in order to reduce the signaling and computation overhead. For each SG, two values of FRF can be applied, either reuse-1 or reuse-N_s, where N_s denotes the number of sectors within one pseudo-cell. When FRF is 1, each sector within the pseudo-cell can utilize all the corresponding sub-carriers in the SG for data transmission. On the other hand, when FRF is N_s, the sub-carriers are divided into N_s sectors, that is, each sector can utilize only the sub-carriers that belong to itself. A two-step approach is implemented, that first distributes the subcarrier groups to the constituent sectors and then arranges each sector to allocate its resource to the users independently as follows:

- First step: the *inter-sector resource allocation* algorithm is used to determine the amount of SGs with FRF = 1 and that with FRF = N_s, and then divide the sub-carriers in SGs with FRF = N_s into the N_s sectors. The inter-sector resource allocation algorithm is adapted to the load condition determined by the number of users in the sector, their geographic distribution, and their requirements. The algorithm works by determining the preferred FRF value of each user. This value is determined by comparing a predetermined threshold to the values of the ratio between the average data rate of a user on all SGs over one super-frame when FRF value is N_s and 1. If the value of the ratio is larger than the threshold, then this means that the interference avoidance gain is large, and accordingly, N_s is taken as the preferred FRF value of this user. Otherwise, the preferred value is set to 1. The computational complexity of this algorithm is $O(M \times (N+1))$, where M is the number of UEs in the sector, and N is the number of SGs.
- Second step: the *intra-sector resource allocation algorithm* is used to allocate the sub-carriers of each sector to the users, which is done independently in each sector. The best possible SG for each user is determined based on the effective data rate achieved by this assignment. To do so, the algorithm first selects the user with the largest ratio of the remaining data rate, i.e., the difference between the requirement and the currently allocated data rate, to the required data rate and then allocate to the selected user a sub-carrier in its best subcarrier group. This process is repeated until no available sub-carrier remains or all user requirements are satisfied. In the case where two or more users do not obtain their required data rate even after completing the above resource allocation process; a reallocation process is performed. In the reallocation process, sub-carriers from the most satisfied users are recalled and then reallocated to those users who did not get the desired data rate. This reallocation process is repeated until the QoS violation ratio

cannot be further reduced. The computational complexity of this algorithm is $O(M \times max(M, N) \times N \times S)$, where M is the number of UEs in the sector, N is the number of SGs, and S is the number of sub-carriers. While this scheme guarantees the required QoS, its computational complexity is relatively high as it depends on the number of sub-carriers and users.

5) $S = T_{eNB}^{CD}(0, 0, f)$ Schemes: In [39], Son *et al.* presented a strategy similar to that proposed in [38] (discussed in Section III.D) but with a reduced complexity. In this scheme, all eNBs can use all the channels for data transmission (reuse-1); however, the original optimization problem that requires intractable computation complexity for global optimality is decomposed into per-eNB sub-problems. At each slot, each eNB needs to determine: (1) which UE is scheduled on each channel, and (2) how much power is allocated for each scheduled UE on each channel. For the eNB to make its decision, only reference users' feedback information is considered instead of considering all other cell's UEs suffering from the eNB interference. A *reference user* refers to the worst UE in another cell receiving the largest interference from the eNB. For an eNB to carry a per-slot resource allocation, a number of parameters for each of the candidate reference users need to be collected by neighboring eNBs and forwarded in advance to that eNB through signaling. The parameters for a reference user are: its weight presented as the inverse of its average throughput, its received signal strength, and its noise plus interference strength. To reduce the per-slot message exchanged between neighboring eNBs, each eNB calculates the average of the needed parameters for all candidate reference users and sends them, infrequently, to its neighboring eNBs. Each eNB maintains a table containing these average values of candidate reference users. Thus, the only thing that needs to be exchanged at each slot is the indices of scheduled users. Once each eNB receives the indices of scheduled users from neighboring eNBs, it uses the parameters in the table corresponding to the received indices. The computational complexity of this scheme is $O(N \times M)$ for UE scheduling and $O(N)$ power allocation, where N is the number of channels and M is the number of UEs. Compared to other coordinated schemes, this scheme shows a relatively low computational complexity. In addition, the use of reference users significantly reduces the the signaling complexity of the scheme.

6) $S = T_{eNB}^{CD}(0, 0, b)$ Schemes: In [40], Kimura *et al.* proposed a distributed coordination scheme for the LTE downlink systems to optimize the fractional frequency reuse (FFR) of the different bands based on the variations in the cell loading. This is done in a distributed manner, with communication between eNBs through the X2 interface. This idea can be viewed as an implementation approach for the concept of dynamic FFR discussed in [41]. In dynamic FFR, cell-centre zone boundaries (across the network) are dynamically adapted (shrink/expand) depending on user behavior, cell load, and interference situation from other neighboring cells. Thus the proposed scheme works on selecting cell-edge bands optimized according to the time-variant environment using the interference level reports, and/or signaling via interface between eNBs. As shown in Figure 16, cell centre areas have
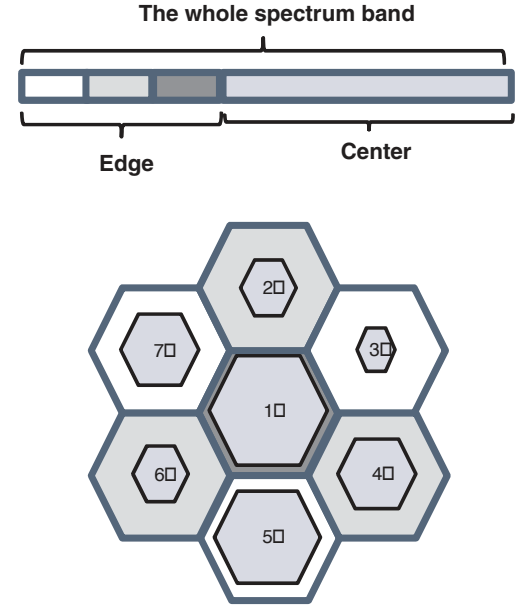

Fig. 16. Concept of dynamic fractional frequency reuse scheme used in [40].

different sizes. For instance, in the figure, Cell 1 is highly loaded while cell 3 experiences a lower load. Accordingly, the region of frequency reuse 1 appears larger in cell 1 compared to cell 3. According to the results reported in [40], this scheme outperforms "reuse-1" and slightly outperforms the static FFR in terms of 5% CDF throughput averaged over the whole time period; while having the same average cell throughput as in the carefully planned static FFR. The scheme adopts static power allocation approach, and thus, all edge-RBs are allocated equal power that is higher than that allocated to center-RBs. This approach reduces the overall computations required to allocate power levels, but as pointed before, it leads to inefficient power distribution.

A limitation in this scheme is that each eNB can only select a pre-determined number of RBs as edge-bands, which limits the capability of the scheme in dealing with *irregular* cell shapes where large number of edge users may exist. This limitation causes the proposed scheme to suffer from the so-called *fake* unavailability of edge-RBs, which may limit the number of served users as in the case of PFR and SFR. Clearly, a crucial design factor in this scheme is the selection of the pre-determined number of edge-RBs. The computation complexity of this scheme is $O(N \times L + N_E)$, where N is the number of all RBs, L is the number of neighbouring eNBs, and N_E is the number of edge RBs. It is worth pointing that the generalized proportional fairness (GPF) scheduling algorithm [51] is adopted in this scheme after defining the edge bands to allocate users to bands, which may increase the overall computational complexity of the algorithm.

E. Autonomous-Distributed Schemes

Similar to coordinated-distributed schemes, resource allocation in autonomous distributed schemes is performed only at the eNB level, with no use of a central entity for coordination. Unlike coordinated-distributed schemes; however, autonomous

distributed schemes require no coordination among eNBs. Instead, each eNB assigns channels to its UEs on its own based only on the local information collected from its UEs. Eliminating the signaling overhead provides the key advantage of autonomous schemes as the performance of the scheduling algorithm becomes independent of the non-negligible latency caused by the interface connecting the eNBs [45]. Thus, decisions can adapt faster to the instantaneous channel and changing traffic conditions. Accordingly, with autonomous schemes, the system becomes *self-organizing* as RBs can be placed anywhere as needed to increase the capacity in a distributed fashion allowing for fast real-time processing. Self-organization is a key factor for the future evolution of mobile networks due to their increasing complexity and required management efforts [61].

To achieve network-wide ICIC and fairness using autonomous schemes, each eNB must altruistically restricts some of its RBs either by reducing their power level or not use them at all, in order to reduce the ICI on these RBs for neighboring eNBs. As there are no communication or coordination between eNBs, autonomous schemes select the RBs that need to be restricted based on the SINR levels on those RBs received by their UE. Clearly, low SINR level indicates that a RB is being used by neighboring eNBs. When restricting RBs, a scheme needs to make a trade-off between the value of reducing the ICI in neighboring cells and the cost of under-utilizing the available spectrum [49].

Due to the complexity of autonomous distributed algorithms, there is a limited, but growing, research effort reported in the literature for developing autonomous distributed ICIC schemes (e.g., [43]- [47]). Results reported for some autonomous distributed schemes show that these schemes can provide a performance comparable to that achieved by various centralized schemes when the number of users is large. It is indeed challenging to devise an ICIC scheme that relies only on the local information, has no control but on the local resources, and yet able to achieve performance similar to that achieved by other coordination schemes discussed above.

In the following, we summarize some of the autonomous-distributed schemes reported in the literature. The schemes are grouped based on the classification dimensions introduced in Section III-A.

1) $S = P_{eNB}^{AD}(0, 0, f)$ Schemes: Stolyar *et al.* proposed a number of schemes in [43] and [44] based on the idea of dynamically allocating power to the different sub-bands to create dynamic soft fractional frequency reuse (SFFR) patterns that can change according to the changes in the workloads and the location of UEs.

In [43], Stolyar *et al.*, presented a self-organizing dynamic SFFR scheme for constant bit-rate (CBR) traffic, which requires no signaling or communication between eNBs. The presented scheme systematically and dynamically achieves a frequency reuse efficient for a given user spatial distribution. The scheme divides the bandwidth into a number of sub-bands, each consists of a number of sub-carriers. Each eNB constantly performs a "selfish" optimization of the assignment of its power and UEs to sub-bands with the objective of optimizing its own performance by minimizing its power usage. This optimization is done based on the other-cell

interference levels reported by UEs for different sub-bands. To understand the idea of self-organizing dynamic SFFR scheme, it is important to observe that, in a given cell: (1) edge UEs generally have larger power requirements, and more importantly, power requirements will be relatively smaller in some "good" sub-bands, where neighboring cells happen to allocate less power, and (2) center UEs generally have smaller power requirements; and they are less affected by the interference from neighboring cells. Accordingly, an eNB that wants to minimize its power usage, it will have a tendency to put its edge UEs into its good sub-bands. Therefore, the cell will allocate larger powers to its good sub-bands; making them become "bad" for the neighboring cells. Neighboring cells then (while trying to minimize their own total powers, as well) will normally "avoid" assigning their edge UEs into those sub-bands, making them even "better" for the cell under consideration, and so on. The system "settles" into a user-to-sub-band allocation pattern, generally requiring less power in all cells, because neighboring cells will automatically "separate" their edge users into different sub-bands. While the scheme results in all sub-bands being utilized in all cells, a self-organizing reuse pattern is created through non-uniform transmission of power across the different sub-bands, as most of the power is transmitted on a subset of the sub-bands, while a small portion of the power is transmitted on the remaining sub-bands. A shadow scheduling algorithm that represents a special case of the Greedy Primal-Dual (GPD) algorithm is used for UEs scheduling, whereas a "fluid" model is used for power allocation of all users in all sub-bands in a cell based on the SINR target and the transmit power that needs to be assigned. A disadvantage of the shadow algorithm is that it solves the problem assuming the sub-carriers and power required by an UE are constant in order to be able to iteratively update the optimal solution instead of keep solving the linear problem from scratch. However; in practice the sub-carriers and power required by an UE may change over time, especially with highly mobile UEs. The computation complexity of this scheme is $O(2 \times M \times N)$, where M is the number of UEs, and N is the number of sub-bands.

In [44], Stolyar *et al.*, presented a power allocation algorithm for automatically adjusting the transmit powers in each sub-band, creating an efficient and dynamic soft fractional frequency reuse (SFFR) patterns for enhancing the performance of OFDMA downlink under *best-effort* traffic. Each eNB runs the algorithm independently and requires no exchange of information with other eNBs. The algorithm runs a fixed number of *virtual time slots* within each physical time slot to assign virtual power allocations to sub-bands until reaching the best allocation. The greater the number of virtual time slots is, the greater the accuracy of the algorithm and its responsiveness to changes in system state, but at an increase computational cost. The algorithm works on assigning power to the different defined sub-bands to achieve a soft frequency reuse of the bands dynamically based on the location of the existing workloads. After power allocation, a generalized proportional fair scheduling algorithm [51] is used for users scheduling. Even though no apriori frequency planning is used, the algorithm; however, improves the cell edge data throughputs, while maintaining the overall cell throughput

at the same level as that achieved by the static approach. The scheme is shown to efficiently achieve a frequency reuse adapted to the UE loads and distribution. The computation complexity of this scheme is $O(N \times V)$, where N is the number of sub-bands, and V is the number of virtual time slots within each physical time slot. It is worth pointing that, this scheme does not consider user scheduling, so adding a user scheduling algorithm will increase the overall computational complexity of the scheme.

A main disadvantage of two schemes proposed by Stolyar *et al.* in [43] and [44] is that, in order to reduce the computation cost, the schemes operate at the granularity of the sub-bands not the RBs. Power allocation at the sub-band granularity level may lead to over allocating power to some of the RBs if they are located in a sub-band that serves edge UEs.

2) $S = P_{UE}^{AD}(r_1, r_2, f)$ *Schemes:* In [45], a distributed resource allocation algorithm is proposed by Cicalo *et al.* to preserve intra-cell fairness and coordinate offline with neighboring cells for load balancing to support inter-cell fairness. The scheme provides an inter-cell interference partial coordination and uses the FFR static scheme for power planning to reduce computational complexity. The authors show that, distributed schemes with aggressive reuse manage to approach the capacity of a centralized system when the number of users is large. One of the disadvantages of this scheme is that it is not fully dynamic as it still requires some static configuration in the FFR power planning that cannot be changed with the change of workloads, and thus, it makes the scheme unsuitable for use in systems with irregular cell configuration. Moreover, in this scheme, the fairness of rate allocation is confined within each cell. To provide global fairness for all users in all cells, an off-line algorithm is used to balances the eNBs load, making the scheme not capable to cope with fast changing workloads and high user mobility. The computation complexity of this scheme is $O(M^2 \times S)$, where M is the number of UEs, and S is the number of subcarriers in the cell. Accordingly, the scheme is dependent on the square of the number of users as well as the number of sub-carriers in the cell. This may limit the applicability of this scheme in LTE-Adv system with 1200 subcarrier at 20MHz [50] and a large number of highly mobile UEs.

3) $S = P_{UE}^{AD}(r_1, r_2, f)$ *Schemes:* In [46], Ko *et al.* introduced a power allocation algorithm called distributed inter-cell power allocation (DIPA) that improves the throughput-fairness trade-off by managing the power allocation to users suffering from high interference without inter-eNB coordination. DIPA is designed to determine the power allocation preferred by each UE individually by using a simplified version of the iterative water-filling scheme, and then incorporate the individual power allocations into one policy. For the iterative water-filling process to converge within the coherence time of the channel, the process is performed over a large time scale (super-frames of 1 sec) based on long-term average channel condition, not adapting to short-term instantaneous channel condition. Each UE measures the average inter-cell interference level on each channel over the previous super-frame period and compares it to the interference threshold broadcasted by the eNB. An UE that suffers from high interference elects a number of channels where the average

interference levels are the lowest, and sends the preference information to the eNB to which it associates. Based on this feedback information, each eNB performs the power allocation procedure that incorporates the preferences of all the users. The power allocation procedure consists of two steps; the first step focuses on the high-interference users (HIU), whereas the second step focuses on the low-interference users. The DIPA algorithm does not consider a specific policy for user scheduling, though proportional fair scheduling algorithm [51] was used in the algorithm evaluation. In contrast to power allocation, user scheduling is performed over a short time scale (frame level of 10 ms) based on the instantaneous channel conditions. This paper; however, did not provide a study to guide the selection of the a proper value of the interference threshold. This value is crucial for operation of the algorithm as it has a clear impact on the balancing the tradeoff between throughput and fairness; as the interference threshold increases so does the fairness at the expensive of a reduced throughput. The computation complexity of this algorithm is $O((N_{HIU} \times M_{HIU})^2 + (N - N_{HIU}) \times (M - M_{HIU}))$, where N_{HIU} is the set of channels preferred by the largest number of users within the high-interference users (HIU) group, M_{HIU} is the number of users in the HIU group, N is the set of all channels in the cell, and M is the set of all users in the cell. The DIPA algorithm does not consider a specific policy for user scheduling, so the DIPA algorithm can be combined with any kind of opportunistic scheduling algorithms, which clearly will add to the computational complexity of the algorithm.

4) $S = I_{UE}^{AD}(0, 0, f)$ *Schemes:* Duy La *et al.* proposed in [47] to use the game theory to solve the resources allocation problem in an autonomous distributed fashion. In this study, UEs (players) compete for a given number of subcarriers by iteratively playing the best-response game till convergence to Nash equilibrium. Each UE works on minimizing the utility function which is defined as the total interferences generated by this UE to the environment plus the total interferences it receives from the environment. The weakness of the formulation presented in [47] is that some strategies may dominate others. For instance, UEs tend to use only single subcarrier to minimize interference leading to inefficient spatial reuse of frequencies. To prevent this dominating strategy, authors forced each UE to be allocated a predetermined fixed number of sub-carriers limiting the OFDMA advantage of allowing UEs to intelligently select favorable sub-carriers for their transmission. Having a fixed number of sub-carriers predetermined for each UE constrains the UEs to a maximum achievable rate. The authors did not provide an algorithm for finding the optimal fixed number of sub-carriers to be allocated to each UE. The computation complexity of this scheme is $O(I \times M \times S!/S_i!(S - S_i)!)$, where I is the expected number of rounds of play until convergence, M is the number of UEs, S is the number of sub-carriers in the system, and S_i is the number of sub-carriers to be allocated to the ith user. This makes the scheme dependent on the number of users as well as the number of sub-carriers in the system. When the total number of sub-carriers is considerably large as the case in LTE-Adv (1200 subcarrier at 20MHz [50]), the convergence rate becomes impractical. Accordingly, the work proposed in [47] assumed unrealistic LTE-Adv environment

of 5 subcarriers for fast convergence given that the minimum supported number of sub-carriers in LTE-Adv is 72 subcarrier at 1.4 Mhz [50]. Finally, the study only focused on the assignment of sub-carriers and left the optimal allocation of power for a further study.

F. Design Trade-offs in Cell Coordination-Based ICIC Schemes

In this subsection, the impact of the different design aspects that affect the performance of the dynamic ICIC schemes is discussed based on the observations obtained from the schemes reviewed above.

- *Coordination*: To perform network-wide ICIC in a multi-cell environment, decisions should be taken in a centralized way at some central coordinator that is able to collect information from all users, and decide accordingly [52]. For a centralized controller to operate, large amount of UEs information need to be exchanged on the backhaul link [30]. Centralized scheduling in multi-cell environment is a difficult task without the use of an efficient and fast infrastructure, due to the stringent time required to exchange the scheduling information and the large feedback required by the UEs [52]. Thus, in practice, the rate of exchanging information with the central coordinators should be kept at low frequencies, which in return may degrade the system's performance. Also, as the number of cells grows, so does the signaling complexity. For these reasons, emerging cellular networks, such as LTE-Advanced systems, rely mainly on inter-eNB coordination over the X2 interface with no central coordinator in a flat architecture [50]. Based on the above discussion, it can be seen that a fully distributed approach with no central controller seems to be more practical in order to facilitate the decision-making process with a small or no signaling delays, which can cope with the instantaneous network conditions. Moreover, eliminating the central controller also reduces the overall infrastructure complexity, and hence, cost. After all, a decision must be made to balance the trade-off between the high eNBs coordination that may increase the overall system throughput, and the high cost of signaling that allows for exchanging sufficient information for coordination [49].

- *Computational complexity*: The problem of resource allocation with dynamic demand is known to be NP-hard [31]. Using an exact method is computationally infeasible as the problem involves extremely large search spaces with correspondingly large number of potential solutions. As the new resource configurations must be computed at run-time, computational efficiency might be of a higher priority compared to model accuracy. One way to reduce the computational complexity of the various schemes in OFDMA systems, is to perform scheduling at a higher granularity levels. In particular, scheduling at the RBs (e.g., [36]) which are in order of tens [50], rather than scheduling at the sub-carrier (e.g., [47]) which are in order of hundreds [50]. The granularity level should be selected carefully to consider the natural reverse relation between granularity and control. For instance, once can schedule sub-bands to user groups as in [43]. This may reduce the overall complexity

even more than in the case of RBs as the number of sub-bands normally is less than ten; however, obviously, this reduction in complexity comes with the cost of limited control in managing the available resources.

- *Signaling complexity*: For all cell-coordination based ICIC schemes, except the autonomous schemes, cooperation between eNBs is required in order to achieve network-wide ICIC in the multi-cell environment. To achieve such cooperation, information need to be exchanged between the eNBs. In practical systems, the LTE X2 interface has a considerable latency [45], resulting in an additional delay on the exchanged information, that should be taken into account, especially in very fast fading environments. Thus, ICIC schemes should minimize the amount of information need to be exchanges in order to comply with the stringent time required to exchange the scheduling information. Schemes that are based on exchanging only the reference UE information (e.g., [39]) instead of all UEs information may result in a considerable reduction in the signaling complexity of the scheme. Another approach to reduce signaling complexity is to exchange only the information of the UEs suffering from interference as in scheme proposed in [38]. Similar to the coordination point above, schemes need to make a trade-off between high coordination between eNBs which may increase the overall system throughput, and the expense of the signaling complexity needed to exchange the required information between eNBs [49].

- *Scheduling rate*: Higher scheduling rates (i.e., more frequent RB allocation updates) bring higher throughput as it allows capturing more accurately the instantaneous channel gain, flat fading, and fast fading [30]. This in turn enables better decisions, especially in the emerging systems that support high user mobility. However, the rate of UE reassignment to different RBs should be kept to minimum to reduce the incurred additional signaling overhead [44]. In dynamic schemes other than autonomous schemes, scheduling is highly dependent on the information exchanged between the eNBs in order to achieve network-wide ICIC in the multi-cell environment. Thus, the scheduling rate is highly dependent on the latency and reliability of the interfaces connecting the eNBs along with the amount of information need to be exchanged between them.

- *Fairness*: To guarantee fairness to edge UEs, a minimum target rate can be pre-defined and fixed for all edge UEs as in the scheme proposed in [33]. This minimum rate constraint forces the rate of each edge UE to be at least as large as this minimum rate. Similar minimum rate can be set for center UEs to guarantee fairness among center UEs as well. A scheme needs to make a trade-off between achieving higher overall throughput and being fair in allocating resources to all UEs [49].

- *Optimization of edge-UEs power allocation*: In traditional frequency reuse schemes, different disjoint sub-channel subsets are assigned to different cells, with sub-channel subsets reused at spatially separated locations. This concept exploits the fact that since the signal power falls off with distance, the same frequency spectrum can be reused at spatially-separated locations. As edge UEs suffer from low SINR on all RBs due to the presence of ICIC and significant path-

loss; thus, the most efficient way to increase their rate is to allocate more power to these UEs instead of more bandwidth [33]. However, as the RBs used by edge UEs might be also used in neighboring cells, cooperation between all cells is required to reduce the ICI resulting for using high power for the edge UEs. In general, manipulating the RB allocated power can achieve higher spectral efficiency. This is done as the same RB is used in different cells, but at different power levels. With power control, ICI can be reduced by reducing the power levels of the dominating interferers.

- *Optimization of center-UEs RB allocation*: While edge UEs suffer the most, and thus, they were typically the main or even only focus of most proposed dynamic schemes, maximizing the rate of center UEs leads to a higher system utilization. Center UEs generally have high SINR on most RBs since they are closer to the serving eNB and far from the interfering eNBs. Thus, the most effective way to improve the rate of these UEs is to allocate more bandwidth instead of power [33].

- *Initialization with pre-partitioned resources between cells and/or user groups*: Using a static frequency allocation scheme to pre-plan the frequency and/or power allocation among cells and/or among user groups can help in reducing the computational complexity of the scheduler as demonstrated in some schemes proposed in the literature (e.g., [33], [40], [44]). However, this planned allocation limits the dynamicity of the scheme. For example, pre-planning the cells bands would limit the scheduler to schedule UE only to RBs in the cell's allocated band, preventing the scheduler from exploring other bands even if this would lead to a better performance and/or improved resources utilization. In case of pre-planned resources partitioning among user groups, the scheduler will not be able to use unused resources reserved for one user group to serve unsatisfied UEs in some other user group that has used its entire pre-allocated band.

IV. FUTURE RESEARCH DIRECTIONS

In this section, we identify and discuss some challenges related to the design and analysis of ICIC scheme that make interesting future research directions.

- *Evaluation Framework and Benchmark*: Due to the complexity of the dynamic ICIC problem, most of the performance evaluations are based on simulation models. A principal problem with simulation evaluations during comparing different schemes is the lack of common context, scenarios and evaluation metrics [48]. Accordingly, an evaluation framework and a benchmark are needed to allow researchers to develop and evaluate their ICIC schemes in a sound manner. Such an effort will provide researchers with data sets for unified realistic scenarios (similar to those provided by the European Momentum project [59] for Berlin and Lisbon) that define common realistic conditions, such as: cells layout, number of channels, propagation data, and traffic intensity. The evaluation framework should provide a unified set of metrics that can be used to evaluate and benchmark various ICIC schemes.

- *Wireless Network Cloud (WNC)*: Recently, with the emergence of the cloud-computing technology and other tech-

nologies related to wireless infrastructure including software radio technology and remote radio head technology, *Wireless Network Cloud (WNC)* with *Base Station Pooling (BSP)* is becoming an interesting alternative network architecture where all eNBs computational resources (enabled by Software Radio) are pooled in a central location and connected via fiber to simple radio-front ends (Remote Radio heads) mounted on remote cell-towers [54]. WNC provides all the necessary transmission and processing resources for a wireless access network to operate in a central fashion [53]. A promising research direction is to re-think the way ICIC centralized schemes are structured by exploiting the transmission and processing resources of the WNC.

- *Heuristics algorithms*: An important line of work is to formulate the ICIC problem as an *optimization task* whose objective is to maximize the multi-cell throughput subject to: power constraints, inter-cell signaling limitations, fairness objectives, and/or minimum bit rate requirements [49]. The problem of resource allocation with dynamic demand is known to be NP-hard [31]. Using an exact method is computationally inefficient as the problem involves extremely large search spaces with correspondingly large number of potential solutions. While optimization models give an insight into the upper bounds of achievable ICIC gains, actually implementing these near optimal mechanisms can be economically and/or technologically infeasible. Thus, various lower complexity heuristics algorithms should be investigated as they have the power of obtaining good suboptimal solutions in a computationally efficient way.

- *Autonomous schemes*: *Static schemes* suffer from the limitation of being unable to adapt to inhomogeneous traffic load and varying user group distribution within each cell [33]. *Centralized and semi-distributed schemes* are often too heavy for implementations in reality as all the interference information on all RBs has to be gathered at a central controller, which is prohibitively large [30]. *Coordinated-distributed schemes* realization has remained limited largely due to constraints on inter-eNB communication and the latencies involved in information exchange for distributed eNBs [53]. *Self-organization* is a key factor for the future evolution of mobile networks, due to their increasing complexity and required management efforts [61]. Thus, with the current network architecture and large number of cells, it appears that the future is for *autonomous schemes* as they can achieve a good ICIC level with no signaling overheads. Moreover, they open the way for a more flexible and adapted cell topology as well as for new energy saving methods [61]. Not much research efforts have been reported in developing autonomous distributed ICIC schemes, which makes it an interesting research direction that is worth further investigation.

V. SUMMARY

Inter-cell interference coordination (ICIC) schemes can be viewed as a scheduling strategy used to limit the inter-cell interference such that cell-edge users in different cells preferably are scheduled on complementary parts of the spectrum when needed. The common theme of ICIC avoidance schemes is to apply restrictions to the usage of downlink resources

such *time/frequency* and/or *transmit power* resources. Such coordination of restrictions will provide an opportunity to limit the interference generation in the area of the cellular network. Accordingly, Signal to Interference and Noise Ratio (SINR) can be improved at the receivers in the coverage area, which will provide potential for increased (cell-edge) data-rates over the coverage area, or increased coverage for given data-rates.

This paper presents a survey of various ICIC schemes used to alleviate the inter-cell interference problem in the downlink of OFDMA-based cellular networks in order to improve cell-edge data rates and enhance the overall network capacity. The paper reviews various ICIC schemes proposed under the two main interference avoidance classes, namely, the *frequency reuse-based* (static) and the *cell coordination-based* (dynamic). To this end, new classification models are presented and used to describe the various schemes under the static and dynamic ICIC avoidance classes.

Static ICIC schemes use methods to allocate subcarriers (frequency bands) among cells and sectors. Whereas dynamic schemes perform real-time cell coordination to allocate resources (frequency bands) to cells and sectors. Pre-allocated frequency reuse schemes are generally simple to implement; however, it is not easy to modify their frequency distributions among adjacent cells in response to the dynamics of the network. Cell coordination schemes, on the other hand, are more flexible as compared to frequency reuse schemes; however, they require complex signaling and coordination among cells.

Generally speaking, frequency reuse schemes are suitable for networks with a static even distribution of loads; however, they lead to significant performance degradation in terms of cell and user throughput when used due to the natural dynamic nature of cellular systems, where there is an unevenly distributed dynamically changing load. Therefore, dynamic frequency allocations are needed in order to cope with the continuous load changes in cells. In cell coordination, interference reduction is realized by real time coordination among all involved cells to avoid that two cell edge UEs in neighboring cells use the same sub-carriers. Adaptive algorithms are developed in order to efficiently manage the resource utilization among cells without a priori resource partitioning. Coordination between cells can be performed in either a centralized, semi- distributed or distributed fashion. Dynamic ICIC schemes reported in the literature are mostly either semi-distributed or distributed via coordination. A limited number of autonomous distributed ICIC schemes have been proposed, and accordingly, more research efforts are needed in order to develop autonomous schemes that can cope with the nature and needs of the emerging OFDMA-based cellular networks with highly mobile users.

REFERENCES

[1] E-UTRA and E-UTRAN Overall description: Stage 2 (Release 8), 3GPP Technical Specification TS 36.300 V8.70, Dec. 2008. Avaliable (July 2012): http://www.3gpp.org.

[2] Requirements for Further Advancements for E-UTRA (LTE-Advanced) (Relase 8), 3GPP Technical Specification TS 36.913 V8.0.0, Jun. 2008. Avaliable (July 2012): http://www.3gpp.org.

[3] 3GPP R1-050629, "Inter-cell interference mitigation," Huawei, 2005. Available (July 2012): http://www.3gpp.org/ftp/tsg_ran/wg1_rl1/TSGR1_AH/LTE_AH_June-05/Docs/R1-050629.zip

[4] 3GPP R1-050764, "Inter-cell Interference Handling for E-UTRA," Ericsson, 2005. Available (July 2012): http://www.3gpp.org/ftp/tsg_ran/WG1_RL1/TSGR1_42/Docs/R1-050764.zip

[5] 3GPP R1-050841, "Further Analysis of Soft Frequency Reuse Scheme," Huawei, 2005. Available (July 2012): http://www.3gpp.org/ftp/tsg_ran/WG1_RL1/TSGR1_42/Docs/R1-050841.zip

[6] 3GPP R1-060291, "OFDMA Downlink inter-cell interference mitigation," Nokia, 2006. Available (July 2012): http://www.3gpp.org/ftp/tsg_ran/WG1_RL1/TSGR1_44/Docs/R1-060291.zip

[7] 3GPP R1-060368, "Performance of Inter-Cell Interference Mitigation with Semi-Static Frequency Planning for EUTRA Downlink," Texas Instruments, 2006. Available (July 2012): http://www.3gpp.org/ftp/tsg_ran/WG1_RL1/TSGR1_44/Docs/R1-060368.zip

[8] R. Kwan and C. Leung, "A Survey of Scheduling and Interference Mitigation in LTE," Electrical and Computer Engineering, vol. 2010, 2010.

[9] 3GPP R1-050507, "Soft Frequency Reuse Scheme for UTRAN LTE," Huawei, 2005. Available (July 2012): http://www.3gpp.org/ftp/tsg_ran/WG1_RL1/TSGR1_41/Docs/R1-050507.zip

[10] N. Hassan and M. Assaad, "Optimal Fractional Frequency Reuse (FFR) and resource allocation in multiuser OFDMA system," in Proc. Int. Conf. Information and Communication Technologies ICICT09., 2009, pp. 88-92.

[11] F. Khan, LTE for 4G Mobile Broadband: Air Interface Technologies and Performance.: Cambridge University Press, 2009.

[12] M. Rahman and H. Yanikomeroglu, "Enhancing cell-edge performance: a downlink dynamic interference avoidance scheme with inter-cell coordination," IEEE Trans. Wireless Commun., vol. 9, no. 4, pp. 1414-1425, 2010.

[13] M. Porjazoski and B. Popovski, "Analysis of Intercell interference coordination by Fractional frequency reuse in LTE," in Proc. Int Software, Telecommunications and Computer Networks (SoftCOM) Conf., 2010, pp. 160-164.

[14] M. Porjazoski and B. Popovski, "Contribution to analysis of Intercell interference coordination in LTE: A fractional frequency reuse case," in Proc. Global Mobile Congress (GMC)., 2010, pp. 1-4.

[15] L. Chen and D. Yuan, "Soft frequency reuse in large networks with irregular cell pattern: How much gain to expect?," in Proc. IEEE 20th Int Personal, Indoor and Mobile Radio Communications Symp., 2009, pp. 1467-1471.

[16] IEEE C802.16m-08/782, "Fractional Frequency Reuse in Uplink," LG Electronics, 2008. Available (July 2012): www.ieee802.org/16/tgm/contrib/C80216m-08_782.doc

[17] Y. Kwon, O. Lee, J. Lee, and M. Chung, "Power Control for Soft Fractional Frequency Reuse in OFDMA System," in Computational Science and Its Applications ICCSA 2010, David Taniar et al., Eds.: Springer Berlin / Heidelberg, 2010, vol. 6018, ch. 7, pp. 63-71.

[18] E. Yaacoub and Z. Dawy,"A Survey on Uplink ResourceAllocation in OFDMA Wireless Network," IEEE Commun. Surveys & Tutorials, VOL. 14, NO. 2, SECOND QUARTER 2012, pp. 322-337, 2012.

[19] D. Kim, J. Ahn, and H. Kim, "Downlink Transmit Power Allocation in Soft Fractional Frequency Reuse Systems," ETRI Journal, vol. 33, pp. 1-5, 2011.

[20] R. Ghaffar and R. Knopp, "Fractional frequency reuse and interference suppression for OFDMA networks," in Proc. 8th Int Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt) Symp., 2010, pp. 273-277.

[21] F. B. Mugdim, "Interference Avoidance Concepts," WINNER II project, 2007.

[22] E. Haro, S. Ruiz, D. Gonzalez, M. Garcia-Lozano, and J. Olmos, "Comparison of Different Distributed Scheduling Strategies for Static/Dynamic LTE Scenarios," Technical University of Wien, 2009.

[23] K. T. Kim and S. K. Oh, "An Incremental Frequency Reuse Scheme for an OFDMA Cellular System and Its Performance," in Proc. IEEE Vehicular Technology Conf. VTC Spring 2008., 2008, pp. 1504-1508.

[24] Z. Xie and B. Walke, "Resource allocation and reuse for inter-cell interference mitigation in OFDMA based communication networks," in Proc. 5th Annual ICST Wireless Internet Conf. (WICON)., 2010, pp. 1-6.

[25] C.-S. Chiu and C.-C. Huang, "Combined Partial Reuse and Soft Handover in OFDMA Downlink Transmission," in Proc. IEEE Vehicular Technology Conf. VTC Spring 2008., 2008, pp. 1707-1711.

[26] A. Gjendemsjo, D. Gesbert, G. E. Oien, and S. G. Kiani, "Optimal Power Allocation and Scheduling for Two-Cell Capacity Maximization," in Proc. 4th Int Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks Symp., 2006, pp. 1-6.

[27] Suman Das, Harish Viswanathan, and G. Rittenhouse, "Dynamic load balancing through coordinated scheduling in packet data systems," in Proc. IEEE INFOCOM 2003., 2003, vol. 1, pp. 786-796.

[28] Thomas Bonald, Sem Borst, and Alexandre Proutiere, "Inter-Cell Scheduling in Wireless Data Networks," in Proc. 11th European Wireless Conference 2005., 2005, pp. 1-7.

[29] M. Rahman and H. Yanikomeroglu, "Multicell Downlink OFDM Subchannel Allocations Using Dynamic Intercell Coordination," in Proc. IEEE Global Telecommunications Conf. GLOBECOM '07., 2007, pp. 5220-5225.

[30] G. Li and H. Liu, "Downlink Radio Resource Allocation for Multi-Cell OFDMA System," IEEE Trans. Wireless Communications, vol. 5, no. 12, pp. 3451-3459, 2006.

[31] M.C. Necker, "A Novel Algorithm for Distributed Dynamic Interference Coordination in Cellular Networks," in Proc. KiVS, pp. 233-238, 2011.

[32] A. Triki and L. Nuaymi, "Intercell Interference Coordination Algorithms in OFDMA Wireless Systems," in Proc. IEEE 73rd Vehicular Technology Conf. (VTC Spring)., 2011, pp. 1-6.

[33] T. Q.S. Quek, Z. Lei, and S. Sun, "Adaptive interference coordination in multi-cell OFDMA systems," in IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications, 2009, pp. 2380-2384.

[34] K. Dong, H. Tian, X. Li, and Q. Sun, "A Distributed Inter-Cell Interference Coordination Scheme in Downlink Multicell OFDMA Systems," in Proc. 7th IEEE Consumer Communications and Networking Conf. (CCNC)., 2010, pp. 1-5.

[35] H. Kwon, W.-I. Lee, and B. Gi Lee, "Low-Overhead Resource Allocation with Load Balancing in Multi-cell OFDMA Systems," in IEEE 61st Vehicular Technology Conference, 2005, pp. 3063-3067.

[36] I. G. Fraimis, V. D. Papoutsis, and S. A. Kotsopoulos, "A distributed radio resource allocation algorithm with interference coordination for multi-cell OFDMA systems," in Proc. IEEE 21st Int Personal Indoor and Mobile Radio Communications (PIMRC) Symp., 2010, pp. 1354-1359.

[37] I. G. Fraimis, V. D. Papoutsis, and S. A. Kotsopoulos, "A Decentralized Subchannel Allocation Scheme with Inter-Cell Interference Coordination (ICIC) for Multi-Cell OFDMA Systems," in Proc. IEEE Global Telecommunications Conf. GLOBECOM 2010., 2010, pp. 1-5.

[38] S. Zheng, H. Tian, Z. Hu, L. Chen, and J. Zhu, "QoS-Guaranteed Radio Resource Allocation with Distributed Inter-Cell Interference Coordination for Multi-Cell OFDMA Systems," in Proc. IEEE 71st Vehicular Technology Conf. (VTC 2010-Spring)., 2010, pp. 1-5.

[39] K. Son, S. Lee, Y. Yi, and S. Chong, "Practical dynamic interference management in multi-carrier multi-cell wireless networks: A reference user based approach," in 8th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, 2010, pp. 186 - 195.

[40] D. Kimura, Y. Harada, and H. Seki, "De-Centralized Dynamic ICIC Using X2 Interfaces for Downlink LTE Systems," in Proc. IEEE 73rd Vehicular Technology Conf. (VTC Spring)., 2011, pp. 1-5.

[41] S.-E. Elayoubi, O. Ben Haddada, and B. Fourestie, "Performance evaluation of frequency planning schemes in OFDMA-based networks," IEEE Trans. Wireless Commun., vol. 7, no. 5, pp. 1623-1633, 2008.

[42] S. Sadr, A. Anpalagan and K. Raahemifar,"Radio Resource Allocation Algorithms for the Downlink of Multiuser OFDM Communication Systems," IEEE Commun. Surveys & Tutorials, vol. 11, no. 3, THIRD QUARTER 2009, pp. 92-106, 2009.

[43] A. L. Stolyar and H. Viswanathan, "Self-Organizing Dynamic Fractional Frequency Reuse in OFDMA Systems," in Proc. INFOCOM 2008. The 27th Conf. Computer Communications. IEEE., 2008, pp. 691-699.

[44] A. L. Stolyar and H. Viswanathan, "Self-Organizing Dynamic Fractional Frequency Reuse for Best-Effort Traffic through Distributed Inter-Cell Coordination," in Proc. IEEE INFOCOM 2009., 2009, pp. 1287-1295.

[45] S. Cicalo, V. Tralli, and A.I. Perez-Neira, "Centralized vs Distributed Resource Allocation in Multi-Cell OFDMA Systems," in IEEE 73rd Vehicular Technology Conference, 2011, pp. 1 - 6.

[46] S. Ko, H. Seo, H. Kwon, and B. G. Lee, "Distributed Power Allocation for Efficient Inter-cell Interference Management in Multi-cell OFDMA Systems," in 16th Asia-Pacific Conference on Communications, 2010, pp. 243-248.

[47] Q. D. La, Y. H. Chew, and B.-H. Soong, "An Interference Minimization Game Theoretic Subcarrier Allocation Algorithm for OFDMA-Based Distributed Systems," in IEEE Global Telecommunications Conference, 2009, pp. 1-6.

[48] I. Katzela and M. Naghshineh, "Channel assignment schemes for cellular mobile telecommunication systems: a comprehensive survey," IEEE Personal Commun., vol. 3, no. 3, pp. 10-31, 1996.

[49] G. Fodor, Ch. Koutsimanis, A. Rcz, N. Reider, A. Simonsson, and W. Mller, "Intercell Interference Coordination in OFDMA Networks and in the 3GPP Long Term Evolution System," Journal of Communications, vol. 4, no. 7, pp. 445-453, 2009.

[50] S. Sesia, I. Toufik, and M. Baker, "LTE — The UMTS Long Term Evolution: from theory to practice," Wiley publishing, 2009.

[51] C. Wengerter, J. Ohlhorst and A. G. E. von Elbwart, "Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA," in Proc. IEEE 61st Vehicular Technology Conf. VTC Spring 2005., 2005, vol.3, pp.1903-1907.

[52] V. Corvino, D. Gesbert and R. Verdone, "A Novel Distributed Interference Mitigation Technique using Power Planning," in Proc. IEEE Wireless Communications and Networking Conf. WCNC 2009, 2009, pp.1-6.

[53] P. Gupta, A. Vishwanath, S. Kalyanaraman and Y. H. Lin, "Unlocking Wireless Performance with Co-operation in Co-located Base Station Pools," In Proc. Second International Conference on Communication Systems and Networks (COMSNETS), 2010, pp. 1-8.

[54] Z. Zhu, Q. Wang, Y. Lin, P. Gupta, S. Sarangi, Sh. Kalyanaraman and H. Franke, "Virtual Base Station Pool: Towards A Wireless Network Cloud for Radio Access Networks," In Proc. 8th ACM International Conference on Computing Frontiers (CF '11), 2011, pp.34:1-34:10.

[55] D. Gonzalez, M. Garcia-Lozano, S. Ruiz and J. Olmos, "Static Inter-Cell Interference Coordination Techniques for LTE Networks: A Fair Performance Assessment," In Proc. the Third international conference on Multiple access communications (MACOM'10), 2010, pp.211-222.

[56] D. Gonzalez, M. Garcia-Lozano, S. Ruiz and J. Olmos, "On the Performance of Static Inter-cell Interference Coordination in Realistic Cellular Layouts," Mobile Networks and Management (MONAMI 2010), Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering (LNICST), vol. 68, no. IV, pp.163-176, 2011.

[57] S. Zhang, "Inter-cell interference coordination in indoor LTE systems," Masters Thesis, Royal Institute of Technology, Sweden, 2011.

[58] R. Verdone, H. Buehler, N. Cardona, A. Munna, R. Patelli, S. Ruiz, P. Grazioso, A. Zanella, A. Eisenblatter and H. Geerdes, "MORANS White Paper - Update," Technical Report available as TD(04)062, COST 273, Athens, Greece, pp. 26-28, January 2008.

[59] European project IST-2000-28088, http://momentum.zib.de, 2005.

[60] M. Kamel and K. Elsayed, "Performance Evaluation of a Coordinated Time-Domain eICIC Framework based on ABSF in Heterogeneous LTE-Advanced Networks,". in Proc. IEEE Global Telecommunications Conf. (GLOBECOM '12), 2012, (To appear).

[61] U. Barth and E. Kuehn, "Self-organization in 4G mobile networks: Motivation and vision," in Proc. 2010 7th International Symposium on Wireless Communication Systems (ISWCS), 2010, pp. 731 -735.

[62] 3GPP TS 36.213 V10.5.0, "E-UTRA Physical layer procedures," 3rd Generation Partnership Project, 2012. Available (July 2012): http://www.3gpp.org/ftp/Specs/archive/36_series/36.213/36213-a50.zip

Abdelbaset S. Hamza received the B.S. (Hons.) degree in Electronics and Communication Engineering from the Institute of Aviation Engineering and Technology, Cairo, Egypt, in 2008. In 2011, he received the M.Sc. degree in Electronics and Communications Engineering from Cairo University, Cairo, Egypt. He is currently studying towards the Ph.D. degree in Computer Science and Computer Engineering department at the University of Nebraska-Lincoln. In addition, he works as a Graduate Teaching Assistant with the department.

Abdelbaset's research interests include the use of intelligent computation techniques to solve the resource allocation problems in various types of systems such as the cognitive radio networks and multi-core multi-processor real time systems. In addition, he previously worked as a Research Assistant in the 4G++ project at Cairo University with focus on spectrum allocation problem in cognitive radio networks. He worked on developing a new capacity boosting techniques in 4th generation wireless systems such as LTE-Advanced. Other research interests include performance analysis of mobile systems over different fading channels, and the optimization problems in WDM networks, regarding the resource allocation, and minimization problems.

Haitham S. Hamza received the B.S. (Honors) and the M.S. degree in Electronics and Communication Engineering from Cairo University, Cairo, Egypt, in 1998 and 2001, respectively. He received the M.S. and the PhD degrees in Computer Science from the University of Nebraska-Lincoln (UNL) in 2002 and 2006, respectively. He is a Fling fellow to the University of Nebraska-Lincoln. Dr. Hamza is a Fling fellow to the University of Nebraska-Lincoln. He is currently an Associate Professor of Information Technology at Cairo University, Egypt.

Dr. Hamza is the director of Advanced Networks Systems and applications Laboratory (ANSR) at Cairo University where he leads a team of 15 researchers working in various types and aspects of computer networks including LTE, WDM, and semantic IoT technologies. He has published more than 60 papers in various international journals and conferences. Dr. Hamza is the recipient of the Best Paper Award in the Optical Symposium of Broadnets in 2005.

He is the recipient of the Graduate Teaching Assistant Award for the Collage of Art and Science in 2005, and the Outstanding Teaching Assistant Award in the Computer Science and Engineering department at UNL for Years 2003 - 2004 and 2004 - 2005. He also received the Graduate Research Assistant Award for the Collage of Art and Science in 2006.

Shady S. Khalifa is currently a PhD candidate at Queen's University, Canada. He received his B.Sc. in 2006 and M.Sc. in 2011 both in Computer Engineering (Networks) from the Dept. of Information Technology, Faculty of Computers and Information, Cairo University, Egypt.

Shady is a member of the ANSR Research Lab, 4G++ Research Group, the Department of Network Administration (DNA) Research Team. He is also one of the co-founders of the Cloud Computing reading group, one of the co-founders of the Teaching and Instructing Center (TIC) for instructors preparation and the EMC Academic Alliance program coordinator at the Faculty of Computers and Information, Cairo University. Shady is the developer of the HSBT NS2 high speed Bluetooth simulation model. Shady is a certified Cisco CCNA instructor, an EMC Information Storage Associate (EMCISA) and an IEEE and ACM student member.

His main research interests include Ultra Large Scale Software, Cloud Computing, Virtualization, Big Data, Information Storage, GPGPU Acceleration, Wireless Networks (ICIC in LTE and LTE-A, Bluetooth, UWB), Network Modeling and Simulation.

Khaled M. F. Elsayed received the Ph.D. in Computer Science and Computer Engineering from North Carolina State University in 1995 and the M.Sc. in Engineering Mathematics and the B.Sc. (honors) in Electrical Engineering from Cairo University in 1990 and 1987 respectively. He is currently a professor of Communication Networks and the head of the Computing Group in the Department of Electronics and Communications Engineering in Cairo University. He is also the PI of the 4G++ and the Web of Objects research projects. Previously, he was the chief technical officer of Embedded Wireless Division in SySDSoft which was acquired by Intel in March 2011. He also worked for Nortel Wireless Systems Engineering as a member of technical staff.

Dr. Elsayed area of research is radio resource management in wireless systems, embedded networked systems, and quality-of-service support in multi-service packet networks. He has about seventy publications in in refereed international journals and conferences. He was an editor for the IEEE Communications Magazine from Nov. 1998 till Dec. 2002. He has also been a member of the technical program committees and session chair for several IEEE, IFIP, and ITC conferences. He is a senior member of IEEE.